Cochrane Library

Cochrane Database of Systematic Reviews

Pulse oximetry screening for critical congenital heart defects (Review)

Plana MN, Zamora J, Suresh G, Fernandez-Pineda L, Thangaratinam S, Ewer AK

Plana MN, Zamora J, Suresh G, Fernandez-Pineda L, Thangaratinam S, Ewer AK.
Pulse oximetry screening for critical congenital heart defects.
Cochrane Database of Systematic Reviews 2018, Issue 3. Art. No.: CD011912.
DOI: 10.1002/14651858.CD011912.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Pulse oximetry screening for critical congenital heart defects

Maria N Plana[1], Javier Zamora[2], Gautham Suresh[3], Luis Fernandez-Pineda[4], Shakila Thangaratinam[5], Andrew K Ewer[6]

[1]Clinical Biostatistics Unit, Ramón y Cajal Hospital (IRYCIS), CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain. [2]Clinical Biostatistics Unit, Ramon y Cajal Institute for Health Research (IRYCIS), CIBER Epidemiology and Public Health (CIBERESP), Madrid (Spain) and Women's Health Research Unit, Centre for Primary Care and Public Health, Queen Mary University of London, London, UK. [3]Section of Neonatology, Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA. [4]Pediatric Cardiology Unit, Hospital Ramon y Cajal, Madrid, Spain. [5]Centre for Primary Care and Public Health, Women's Health Research Unit, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [6]Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

Contact address: Andrew K Ewer, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham Women's Hospital, Edgbaston, Birmingham, B15 2TT, UK. a.k.ewer@bham.ac.uk.

Editorial group: Cochrane Neonatal Group.
Publication status and date: New, published in Issue 3, 2018.

Citation: Plana MN, Zamora J, Suresh G, Fernandez-Pineda L, Thangaratinam S, Ewer AK. Pulse oximetry screening for critical congenital heart defects. *Cochrane Database of Systematic Reviews* 2018, Issue 3. Art. No.: CD011912. DOI: 10.1002/14651858.CD011912.pub2.

ABSTRACT

Background

Health outcomes are improved when newborn babies with critical congenital heart defects (CCHDs) are detected before acute cardiovascular collapse. The main screening tests used to identify these babies include prenatal ultrasonography and postnatal clinical examination; however, even though both of these methods are available, a significant proportion of babies are still missed. Routine pulse oximetry has been reported as an additional screening test that can potentially improve detection of CCHD.

Objectives

• To determine the diagnostic accuracy of pulse oximetry as a screening method for detection of CCHD in asymptomatic newborn infants

• To assess potential sources of heterogeneity, including:

○ characteristics of the population: inclusion or exclusion of antenatally detected congenital heart defects;

○ timing of testing: < 24 hours versus ≥ 24 hours after birth;

○ site of testing: right hand and foot (pre-ductal and post-ductal) versus foot only (post-ductal);

○ oxygen saturation: functional versus fractional;

○ study design: retrospective versus prospective design, consecutive versus non-consecutive series; and

○ risk of bias for the "flow and timing" domain of QUADAS-2.

Search methods

We searched the Cochrane Central Register of Controlled Trials (CENTRAL; 2017, Issue 2) in the Cochrane Library and the following databases: MEDLINE, Embase, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Health Services Research Projects in Progress (HSRProj), up to March 2017. We searched the reference lists of all included articles and relevant systematic reviews to identify additional studies not found through the electronic search. We applied no language restrictions.

Selection criteria

We selected studies that met predefined criteria for design, population, tests, and outcomes. We included cross-sectional and cohort studies assessing the diagnostic accuracy of pulse oximetry screening for diagnosis of CCHD in term and late preterm asymptomatic newborn infants. We considered all protocols of pulse oximetry screening (eg, different saturation thresholds to define abnormality, post-ductal only or pre-ductal and post-ductal measurements, test timing less than or greater than 24 hours). Reference standards were diagnostic echocardiography (echocardiogram) and clinical follow-up, including postmortem findings, mortality, and congenital anomaly databases.

Data collection and analysis

We extracted accuracy data for the threshold used in primary studies. We explored between-study variability and correlation between indices visually through use of forest and receiver operating characteristic (ROC) plots. We assessed risk of bias in included studies using the QUADAS-2 tool. We used the bivariate model to calculate random-effects pooled sensitivity and specificity values. We investigated sources of heterogeneity using subgroup analyses and meta-regression.

Main results

Twenty-one studies met our inclusion criteria (N = 457,202 participants). Nineteen studies provided data for the primary analysis (oxygen saturation threshold < 95% or ≤ 95%; N = 436,758 participants). The overall sensitivity of pulse oximetry for detection of CCHD was 76.3% (95% confidence interval [CI] 69.5 to 82.0) *(low certainty of the evidence)*. Specificity was 99.9% (95% CI 99.7 to 99.9), with a false-positive rate of 0.14% (95% CI 0.07 to 0.22) *(high certainty of the evidence)*. Summary positive and negative likelihood ratios were 535.6 (95% CI 280.3 to 1023.4) and 0.24 (95% CI 0.18 to 0.31), respectively. These results showed that out of 10,000 apparently healthy late preterm or full-term newborn infants, six will have CCHD (median prevalence in our review). Screening by pulse oximetry will detect five of these infants as having CCHD and will miss one case. In addition, screening by pulse oximetry will falsely identify another 14 infants out of the 10,000 as having suspected CCHD when they do not have it.

The false-positive rate for detection of CCHD was lower when newborn pulse oximetry was performed longer than 24 hours after birth than when it was performed within 24 hours (0.06%, 95% CI 0.03 to 0.13, vs 0.42%, 95% CI 0.20 to 0.89; P = 0.027).

Forest and ROC plots showed greater variability in estimated sensitivity than specificity across studies. We explored heterogeneity by conducting subgroup analyses and meta-regression of inclusion or exclusion of antenatally detected congenital heart defects, timing of testing, and risk of bias for the "flow and timing" domain of QUADAS-2, and we did not find an explanation for the heterogeneity in sensitivity.

Authors' conclusions

Pulse oximetry is a highly specific and moderately sensitive test for detection of CCHD with very low false-positive rates. Current evidence supports the introduction of routine screening for CCHD in asymptomatic newborns before discharge from the well-baby nursery.

PLAIN LANGUAGE SUMMARY

Pulse oximetry for diagnosis of critical congenital heart defects

Review question

We reviewed evidence on the accuracy of pulse oximetry for detection of critical congenital heart defects (CCHDs) in asymptomatic newborn infants.

Background

CCHDs occur in around two in 1000 newborn infants and are a leading cause of infant death. Timely diagnosis is crucial for best outcomes for these babies, but current screening methods may miss up to 50% of affected newborn infants before birth, and those sent home before diagnosis frequently die or endure major morbidity. However, babies with CCHD often have low blood oxygen levels, which can be detected quickly and non-invasively by pulse oximetry, using a sensor placed on the newborn infant's hand or foot. A pulse oximeter is a machine that can measure, non-invasively, the amount of oxygen carried around the body by red blood cells. Oxygen from the lungs is bound to hemoglobin in red blood cells, forming oxyhemoglobin. If oxygen is not bound, de-oxyhemoglobin is formed. In

health, almost all hemoglobin is oxyhemoglobin, and so oxygen saturation (ie, the percentage of hemoglobin that has bound oxygen) is close to 100%. The pulse oximeter measures this by passing light through peripheral blood vessels (eg, a fingertip in an adult, in a hand or foot in a baby). Oxyhemoglobin and de-oxyhemoglobin absorb this light in different ways, and the proportion of light absorbed can be analyzed by software within the oximeter, which then calculates the percentage of hemoglobin saturated with oxygen.

Study characteristics

We searched until March 2017 for evidence on use of pulse oximetry to detect CCHD in newborn infants and found 21 studies. These studies used different thresholds to define a pulse oximetry test as positive. We combined all studies using a threshold around 95% (19 studies with 436,758 newborn infants).

Key results

This review found that for every 10,000 apparently healthy newborn infants screened, around six of them will have CCHD. The pulse oximetry test will correctly identify five of these newborn infants with CCHD (but will miss one case). Newborn infants who are missed could die or experience major morbidity.

For every 10,000 apparently healthy newborn infants screened, 9994 will not have CCHD. The pulse oximetry test will correctly identify 9980 of them (but 14 newborn infants will be investigated for suspected CCHD). Some of these infants may be exposed to unnecessary additional tests and a prolonged hospital stay, but a proportion will have a potentially serious non-cardiac illness.

The number of newborn infants incorrectly investigated for CCHD decreases when pulse oximetry is performed longer than 24 hours after birth.

Certainty of evidence

We judged the included studies to be mainly at low or unclear risk of bias for several of the certainty domains assessed. Some studies used less robust methods to verify negative results. We considered the overall certainty of the evidence as moderate.

BACKGROUND

Congenital heart defects (CHDs) constitute the most common group of congenital malformations, with an incidence of 4 to 10 per 1000 live births (Botto 2001; Lloyd-Jones 2009; Mahle 2009; Wren 2008); they account for more deaths than any other congenital malformation (Heron 2007; Mahle 2009; Office of National Statistics, 2015), and up to 10% of all infant deaths are attributed to them (Abu-Harb 1994; Boneva 2001; Knowles 2005; Lloyd-Jones 2009; Wren 2008). Life-threatening critical CHDs (CCHDs) account for approximately 15% to 25% of all CHDs (Mahle 2009; Wren 2008). Most CCHDs are amenable to treatment, but poor clinical condition at the time of surgery increases mortality and has been shown to result in worse outcomes for conditions such as hypoplastic left heart (Brown 2001; Brown 2006), coarctation of the aorta (Franklin 2002), and transposition of the great arteries (Tworetzky 2001). Early detection of these conditions can reduce the risk of acute cardiovascular collapse and death (Abu-Harb 1994; Mahle 2009).

Most newborns with a CCHD are asymptomatic at birth (Wren 2008); detection before the onset of symptoms usually involves routine screening by antenatal ultrasound scan, as described by Allan 1986 and Bull 1999, and by postnatal clinical examination of the cardiovascular system, as reported by Hall 1999. Unfortunately, both methods have a variable, and often low, detection rate (Abu-Harb 1994a; Carvalho 2002; Chew 2007; Garne 2001; Tegnander 2006; Westin 2006; Wren 1999), and up to 30% of infants born with CCHDs are discharged home before the diagnosis has been established (Abu-Harb 1994; Brown 2006; Mellander 2006; Wren 2008), with reported mortality rates as high as 50% (Chang 2008).

Although antenatal detection rates following screening ultrasonography are improving, average detection of isolated CCHD remains less than 50% (Abu-Harb 1994a; Carvalho 2002; Chew 2007; Garne 2001; Tegnander 2006; Westin 2006; Wren 1999). Clinical examination abnormalities such as murmur and weak pulse are often absent in early postnatal life, and the more common finding of cyanosis (bluish discoloration of the skin due to reduced oxygen in the blood) is frequently clinically undetectable (Mahle 2008; O'Donnell 2007). The fact that most infants with CCHD will have such mild cyanosis has led to the exploration of pulse

oximetry assessment as a possible screening test to identify affected infants (Ewer 2012a; Knowles 2005; Lloyd-Jones 2009).

Following publication of several large test accuracy studies, several countries adopted pulse oximetry screening as routine practice, and many more are considering its introduction (de-Wahl Granelli 2014; Ewer 2014; Kuelling 2009; Mahle 2012; Manzoni 2017).

In addition to test accuracy, studies have demonstrated that pulse oximetry screening is cost-effective (Knowles 2005; Peterson 2013; Roberts 2012), and that it is acceptable to both parents and clinical staff (Narayen 2017; Powell 2013).

The vast majority of babies studied have been screened in a hospital setting - specifically, the well-baby nursery - at low altitude. However, screening has been reported recently in other settings including neonatal units (Iyengar 2014; Suresh 2013), as well as out of hospital settings such as home births - reported by Cawsey 2016; Lhost 2014; and Narayen 2016a - and births at moderate altitude (Han 2013; Wright 2014).

This review does not include settings outside the well-baby nursery.

Target condition being diagnosed

The definition of CCHD is not consistent, and the literature reveals many interpretations (Ewer 2012a). One of the difficulties arises because some conditions (such as coarctation of the aorta and pulmonary stenosis) may or may not predispose to acute collapse, depending on relative severity. For the purposes of this review, we have used a previously described definition of CCHD, that is, "any potentially life-threatening duct-dependent heart lesion from which infants either die or require invasive procedures (surgery or cardiac catheterization) in the first 28 days of life" (Ewer 2012a; Wren 2008). The definition includes all infants with hypoplastic left heart syndrome, pulmonary atresia with intact ventricular septum, simple transposition of the great arteries, or interruption of the aortic arch. In addition, all infants dying or needing surgery or catheter in the first 28 days of life with coarctation of the aorta, aortic valve stenosis, pulmonary valve stenosis, tetralogy of Fallot, pulmonary atresia with ventricular septal defect, or total anomalous pulmonary venous connection are classified as having critical congenital heart defects. This definition offers the advantages that it allows a degree of assessment of severity of certain lesions based on early death or intervention, it is relatively easy to categorize, and it has been used in several test accuracy studies and previous systematic reviews (Ewer 2011; Ewer 2012a; Thangaratinam 2012; Zhao 2014).

Index test(s)

Pulse oximetry measurement of oxygen saturations in an asymptomatic newborn infant can be used to identify CCHD before discharge from hospital. Pulse oximetry is an accurate and well-established test used to quantify hypoxemia (low oxygen levels in the blood) that is rapid, painless, and easy to perform in all patient groups including newborn infants (Ewer 2012a; Ewer 2013; Knowles 2005; Lloyd-Jones 2009; Mahle 2009; Narayen 2016b). Any trained individual can perform pulse oximetry screening, and results can be obtained in approximately five minutes. For infants enrolled in pulse oximetry screening studies, pulse oximetry probes (to measure oxygen saturations) are placed on the foot only (post-ductal) or on the right hand and foot (pre-ductal and post-ductal) (Ewer 2012b; Ewer 2013; Ewer 2016). The index test allows screening to reduce the number of infants discharged from hospital before diagnosis of CCHD, and it can be performed at any point before discharge before or after the clinical examination.

Clinical pathway

Standard screening for CCHD usually includes midtrimester ultrasonography of pregnant women, which includes assessment of fetal cardiac anatomy. If a cardiac defect is suspected when this examination is performed, a detailed fetal echocardiogram may confirm the diagnosis. Most newborn infants also undergo one or more clinical examinations before discharge from the hospital, which include assessment of the cardiovascular system (auscultation of heart sounds, palpation of peripheral pulses). If a cardiac defect is suspected upon completion of either of these screening tests, then a postnatal diagnostic echocardiogram is usually obtained. As described previously, these screening tests have variable, and often low, detection rates.

The population included in this review may or may not have had antenatal screening. All were asymptomatic at the time of pulse oximetry screening.

Alternative test(s)

In addition to the screening tests already described, alternatives such as routine screening fetal echocardiography and postnatal echocardiography have been proposed but are unlikely to be cost-effective (Knowles 2005). This review did not assess the accuracy of existing screening tests (ie, antenatal ultrasonography and physical examination).

Rationale

Hypoxemia, or suboptimal arterial oxygen saturation, is present in most infants with CCHD (Ewer 2012a; Lloyd-Jones 2009; Mahle 2009). Some may have overt cyanosis, but in many, the degree of hypoxaemia may be difficult to discern on clinical examination. Pulse oximetry is a quick, painless, non-invasive, and reliable method used to determine arterial oxygen saturation levels; it has been widely used in many areas of clinical medicine

for over 30 years. The concept of using oxygen saturations as a screen for critical heart defects was first reported more than 15 years ago. Pulse oximetry screening may allow detection of infants who have been missed by other screening methods before they are discharged from hospital, allowing urgent cardiac intervention before the onset of life-threatening cardiorespiratory collapse. Systematic reviews of pulse oximetry screening studies have been published (Thangaratinam 2007; Thangaratinam 2012), and indeed this screening technique is now common practice in the United States and in some European countries. However, this is the only review that includes recent large test accuracy studies, including one reported from a middle-income country - China.

We performed a systematic review of studies assessing the diagnostic accuracy of screening with pulse oximetry (index test) in relation to echocardiography or clinical follow-up (reference standard) for detection of CCHD in asymptomatic newborn infants. It is important to note that we wanted to determine how well a negative pulse oximetry test result rules out a CCHD diagnosis. Several previous reviews have explored this topic (Ewer 2012a; Ewer 2013; Knapp 2010; Knowles 2005; Lloyd-Jones 2009; Narayen 2016b; Thangaratinam 2007).

OBJECTIVES

- To determine the diagnostic accuracy of pulse oximetry as a screening method for detection of CCHD in asymptomatic newborn infants

Secondary objectives

- To assess potential sources of heterogeneity, including:
    - characteristics of the population: inclusion or exclusion of antenatally detected congenital heart defects;
    - timing of testing: < 24 hours versus ≥ 24 hours after birth;
    - site of testing: right hand and foot (pre-ductal and post-ductal) versus foot only (post-ductal);
    - oxygen saturation: functional versus fractional;
    - study design: retrospective versus prospective design, consecutive versus non-consecutive series; and
    - risk of bias for the "flow and timing" domain of QUADAS-2.

METHODS

Criteria for considering studies for this review

Types of studies

We considered inclusion of prospective or retrospective cohort and cross-sectional studies evaluating the diagnostic accuracy of pulse oximetry as a screening method for detection of critical congenital heart defects in asymptomatic newborn infants. A study of diagnostic accuracy should provide sufficient data for construction of the two-by-two table showing the cross-classification of disease status (CCHD) and test outcome (pulse oximetry). We excluded studies if we could not extract true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) values after contacting corresponding authors of primary studies when necessary. We excluded case reports and studies of case-control design.

Participants

We included studies that recruited asymptomatic (with no signs of respiratory or cardiac illness) term or near-term newborns before discharge from hospital.

Index tests

The test under evaluation was pulse oximetry screening to identify low oxygen saturation. We included all protocols of screening (eg, post-ductal [foot] only vs pre-ductal and post-ductal [right hand and foot], different saturation thresholds to define abnormality, different numbers of repeat tests). Criteria for defining a screen as positive or negative in this review were those used by the authors of respective publications.

Target conditions

Critical congenital heart defects as defined above.

Reference standards

Reference standards were diagnostic echocardiography (echocardiogram) and clinical follow-up in the first 28 days of life, including postmortem findings and information from mortality and congenital anomaly databases, to identify patients with false-negative findings.

Search methods for identification of studies

Electronic searches

Information Specialists of the Cochrane Neonatal Review Group performed the searches. Using the strategy described in Appendix 1, they searched the following databases.

- Cochrane Central Register of Controlled Trials (CENTRAL; 2017, Issue 2) in the Cochrane Library.
- MEDLINE via PubMed (1966 to March 2017).

- Embase via Ovid (1980 to March 2017).
- Cumulative Index to Nursing and Allied Health Literature (CINAHL) (1982 to March 2017).

The MEDLINE search strategy included medical subject headings (MeSH) and free text words (see Appendix 1). We adjusted this strategy for use with the other electronic databases. We considered a combination of medical subject headings and text terms to generate three subsets of citations: one subset indexing the index test (pulse oximetry), a second subset indexing the target population (infant-newborn), and a third subset indexing the clinical condition (congenital heart disease). We combined these subsets to generate a set of citations relevant to our research question. We considered both published and unpublished reports for inclusion and excluded studies published in abstract form only. We applied no language restriction to the electronic searches.

Searching other resources

We used the Science Citation Index, accessed via the Institute for Scientific Information (ISI) Web of Science, to retrieve reports citing the studies included in this review. We searched for similar systematic reviews in the Database of Abstracts of Reviews of Effects (DARE) to March 2017, to cross-reference results. We also searched the Health Services Research Projects in Progress (HSRProj) database (http://www.nlm.nih.gov/hsrproj/) (searched on March 20, 2017). We handsearched the reference lists of all relevant primary studies on the topic of our interest to identify cited articles not captured by our electronic searches (up to March 15, 2017). We applied no language restrictions.

Data collection and analysis

Selection of studies

Two review authors (MNP and JZ) independently screened titles and abstracts identified through electronic literature searches to identify potentially eligible studies. First, we excluded those records classified by both review authors as "excluded." Second, we independently assessed the full text of reports classified as "unsure" or "potentially eligible" by applying the selection criteria outlined above in the Criteria for considering studies for this review section. We resolved disagreements through discussion. If finally we reached no consensus, we consulted a third review author (AKE).

Data extraction and management

We used a standardized data extraction form to aid extraction of relevant information and data from each included study. Three review authors (MNP, LFP, and AKE) separately participated in data extraction. MNP and LFP extracted data corresponding to study design, participant details, method of testing, threshold saturation level, and type of oxygen saturation measured, as well as timing of the test and inclusion or exclusion of infants with suspected congenital heart defects after antenatal ultrasound screening in pregnancy, reference tests, and funding. AKE and MNP extracted the following data to reconstruct the two-by-two table: true-positive, false-positive, true-negative, and false-negative values or, if not available, relevant parameters (sensitivity, specificity, or positive and negative predictive values). Two review authors (MNP and JZ) incorporated data and study characteristics into Review Manager 5.3 (RevMan 2014).

Dealing with duplicate publications

We included only once those studies that have been published in duplicate, ensuring that we extracted all relevant data from all publications.

Inconclusive results

Although we did not anticipate uninterpretable results, when we detected these cases, we excluded them from analysis and adequately reported their frequency in tables.

Assessment of methodological quality

Two review authors (MNP and LFP) independently appraised the methodological quality of each included study using the QUADAS-2 tool (Whiting 2011). QUADAS-2 consists of four domains, each requiring a risk of bias categorization of low, high, or unclear risk. The first three domains are also assessed in terms of concerns about applicability (applicability concerns ratings). Each domain comprises a set of signaling questions that should be marked as "yes," "no," or "unclear." We tailored QUADAS-2 for our specific review question by modifying signaling questions accordingly and providing guidance on how to assess risk of bias and applicability concerns ratings (Appendix 2). We resolved disagreements between risk of bias and applicability concern ratings through discussion or by consultation with a third review author (AKE). We summarized our results in the text and in tables and corresponding figures. We decided post hoc to assess the certainty of evidence by using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach (GRADEpro GDT; Hultcrantz 2017; Schunemann 2008).

Statistical analysis and data synthesis

We performed analyses using methods described in Chapter 10 of the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (Macaskill 2010).

We considered pulse oximetry screening as positive if the oxygen saturation level was below the threshold defined in the primary

study, and negative if it was above that threshold. Cross-classification of these test results with those of the reference standard(s) produced the numbers of true positives, false positives, true negatives, and false negatives for each study, based on the ability of pulse oximetry to detect CCHD.

We used data from the two-by-two tables to calculate sensitivity and specificity for individual studies. We present individual study results by plotting the estimates of sensitivity and specificity (and their 95% confidence intervals) in both forest plots and receiver operating characteristic (ROC) scatter plots. We extracted accuracy data for the threshold used in primary studies.

We performed meta-analyses using a bivariate model (Chu 2006; Reitsma 2005). This model accounts for intra-study accuracy variability and inter-study variations in test performance with inclusion of random effects. We analyzed studies sharing the same threshold and obtained summary accuracy estimates (when the number of studies was enough). We present these estimates with a 95% confidence ellipse in the ROC space. We used pooled estimates of sensitivity and specificity to derive positive and negative likelihood ratios that can be used to update the prior probability of having CCHD to a post-test probability of having CCHD after a positive or negative pulse oximetry result. The greater the positive likelihood ratio and the lower the negative likelihood ratio, the more important the effect of the test on changing pretest into post-test probabilities. We did not calculate positive and negative predictive values because these indices depend on the prevalence of the target condition (ie, CCHD).

For analyses, we used a METADAS SAS macro that estimates parameters for the model with SAS Proc NLMIXED (SAS Institute Inc. 2004; Takwoingi 2010). We entered parameter estimates from the bivariate model into RevMan to produce the summary operating point with a 95% confidence region and a 95% prediction region (Chu 2006; Reitsma 2005; RevMan 2014).

Investigations of heterogeneity

We explored between-study variability and correlation between indices visually through forest and ROC plots. We measured total between-study variability in sensitivity and in specificity through variances of the random effects for logit(sensitivity), logit(specificity), and their covariance of the bivariate model. We also provided confidence and prediction ellipses. We further investigated heterogeneity by exploring effects of several study-level factors through subgroup and meta-regression analyses including covariate terms to the bivariate model (Chu 2006; Reitsma 2005). When available, we examined the following covariates.

- Inclusion or exclusion of antenatally detected congenital heart defects.
- Screening test method (the screening test may be performed at different times after birth, oxygen saturations may be measured at pre-ductal and post-ductal sites or at post-ductal sites only, and, finally, the oxygen saturation measured could be expressed as "functional" [which refers to the proportion of oxygenated hemoglobin that is *capable* of binding oxygen] or "fractional" [which refers to the percentage of total hemoglobin that is oxygenated]). In most cases, differences between the two values are very small, and most modern pulse oximeters measure functional saturations only.
- Study design (included studies may be prospective or retrospective, and may enroll consecutive patients or not). We expect that retrospective studies are more prone to information and selection biases. In this review, it is more likely that medical records of infants with a positive index test result include more information as compared with medical records of infants with a negative test result (information bias). In a similar way, infants with any CCHD are more likely than infants without CCHD to be detected and included in the study after a retrospective medical records review (selection bias).
- Risk of bias of the "flow and timing" domain of the QUADAS-2 questionnaire (unclear/high vs low risk of bias). It is expected that studies used different reference standards to confirm index test results (echocardiogram, clinical follow-up, registries in mortality, and congenital anomaly databases).

Sensitivity analyses

We examined the robustness of meta-analyses by conducting sensitivity analyses. We checked the impact of excluding studies from analysis according to domains of the QUADAS-2 assessment. Additionally, we decided to perform ad hoc sensitivity analyses to explore how sensitivity and specificity vary by including or excluding studies with different thresholds.

Assessment of reporting bias

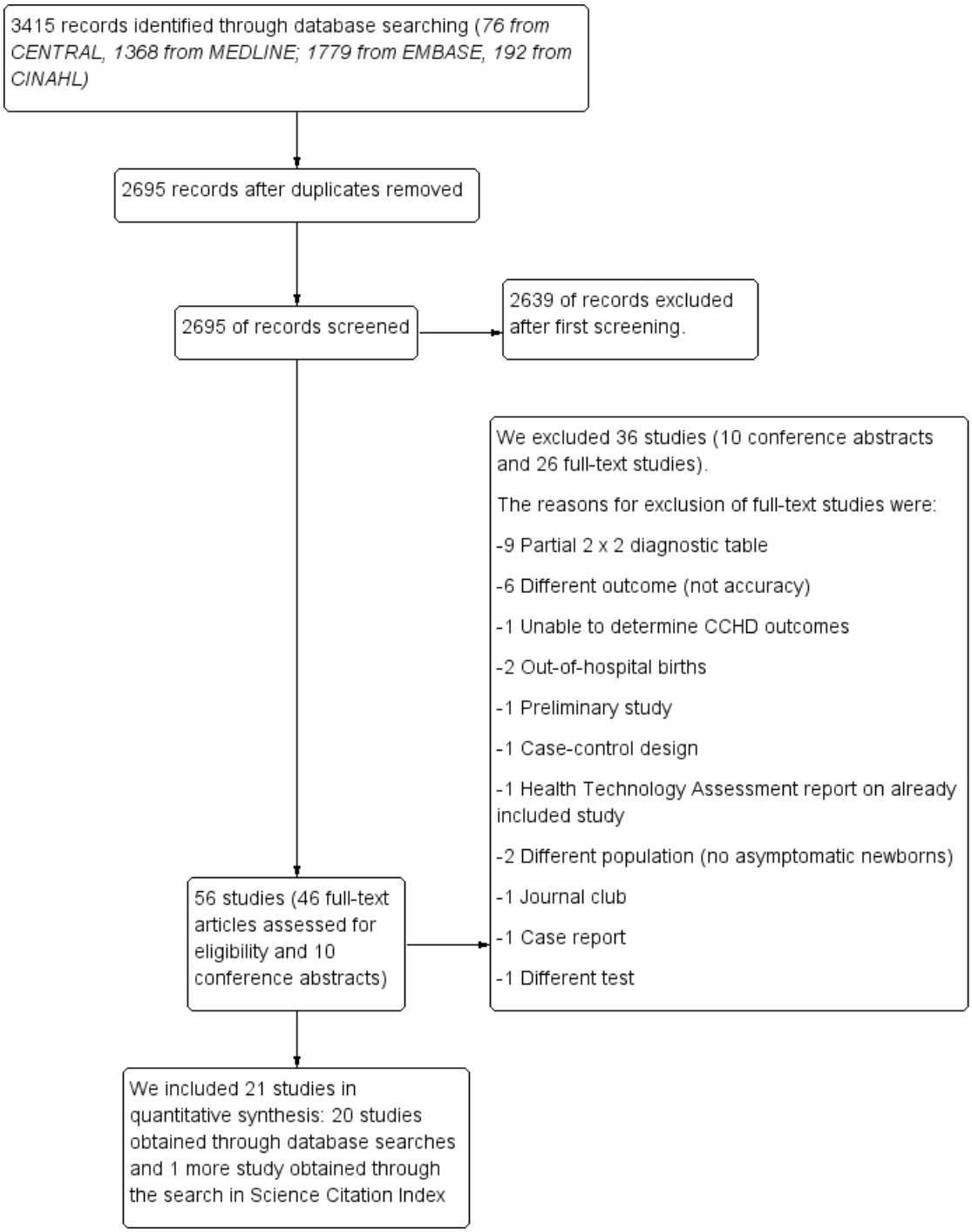
We did not investigate reporting bias, given the limited power of available tests and uncertainty about interpreting statistical evidence of funnel plot asymmetry as necessarily implying publication bias (Leeflang 2008).

RESULTS

Results of the search

Figure 1 shows details of the search and selection process. Electronic database searches yielded a total of 3415 references from CENTRAL, MEDLINE, Embase, and CINAHL. Searches for primary studies through other resources did not reveal additional potentially eligible studies.

Figure 1. Flow of studies through the screening process. CCHD: critical congenital heart defect.



After de-duplication, two review authors (MNP and JZ) independently assessed 2695 references against the inclusion criteria. During initial screening of titles and abstracts, we identified 56 studies (46 full-text papers and 10 conference abstracts). We excluded 2639 references because they did not meet the inclusion criteria. We also excluded those published in abstract form only (n = 10). Of 46 full-text studies, nine studies provided a partial two-by-two diagnostic table, and we excluded them. We excluded 17 other studies for the following reasons.

- Different outcomes (not accuracy) (n = 6).
- Inability to determine CCHD outcomes (n = 1).
- Out-of-hospital births (n = 2).
- Preliminary studies (n = 1).
- Different population (n = 2).
- Health technology assessment report on already included study (n = 1).
- Case-control study (n = 1).
- Journal club (n = 1).
- Case report (n = 1).
- Different index test (n = 1) (see Characteristics of excluded studies).

We obtained one additional study by searching Science Citation Index (Gomez-Rodriguez 2015). We included 21 studies in a quantitative synthesis (Arlettaz 2006; Bakr 2005; Bhola 2014; de-Wahl Granelli 2009; Ewer 2011; Gomez-Rodriguez 2015; Jones 2016; Klausner 2017; Koppel 2003; Meberg 2008; Oakley 2015; Ozalkaya 2016; Richmond 2002; Riede 2010; Rosati 2005; Sendelbach 2008; Singh 2014; Turska 2012; Van Niekerk 2016; Zhao 2014; Zuppa 2015).

Characteristics of studies

We provide in the Characteristics of included studies table details on the design, setting, population, index test, target condition, and reference standard of all included studies. We prepared an additional table (Table 1) to summarize the main characteristics. Of 3415 references, we identified 21 primary studies that were eligible for inclusion and provided data for 457,202 newborn infants (Figure 1). Studies were published between 2002 and 2017. Countries included were United Kingdom (Ewer 2011; Jones 2016; Oakley 2015; Richmond 2002; Singh 2014), Italy (Rosati 2005; Zuppa 2015), USA (Klausner 2017; Koppel 2003; Sendelbach 2008), Australia (Bhola 2014), China (Zhao 2014), Germany (Riede 2010), Mexico (Gomez-Rodriguez 2015), Norway (Meberg 2008), Poland (Turska 2012), Saudi Arabia (Bakr 2005), South Africa (Van Niekerk 2016), Sweden (de-Wahl Granelli 2009), Switzerland (Arlettaz 2006), and Turkey (Ozalkaya 2016).

Sixteen studies included prospective cohorts (Arlettaz 2006; Bakr 2005; de-Wahl Granelli 2009; Ewer 2011; Gomez-Rodriguez 2015; Koppel 2003; Meberg 2008; Oakley 2015; Richmond 2002; Riede 2010; Rosati 2005; Sendelbach 2008; Turska 2012; Van Niekerk 2016; Zhao 2014; Zuppa 2015), as well as five retrospective cohorts (Bhola 2014; Jones 2016; Klausner 2017; Ozalkaya 2016; Singh 2014). Seventeen studies excluded newborns who were suspected to have congenital heart disease after antenatal ultrasound screening during pregnancy (Bakr 2005; de-Wahl Granelli 2009; Gomez-Rodriguez 2015; Jones 2016; Klausner 2017; Koppel 2003; Meberg 2008; Oakley 2015; Ozalkaya 2016; Riede 2010; Rosati 2005; Sendelbach 2008; Singh 2014; Turska 2012; Van Niekerk 2016; Zhao 2014; Zuppa 2015) (Table 1).

Nine studies performed pulse oximetry within 24 hours after birth (Arlettaz 2006; Ewer 2011; Gomez-Rodriguez 2015; Jones 2016; Meberg 2008; Richmond 2002; Sendelbach 2008; Singh 2014; Turska 2012) (Table 1). Twelve studies used the foot alone (post-ductal) to measure oxygen saturation, and the remainder used both right hand and foot (pre-ductal and post-ductal) (Table 1). Investigators used several different pulse oximeter models (see description in Table 1). Two studies measured fractional saturations (Bakr 2005; Richmond 2002) (Table 1). Eight studies used a post-ductal saturation threshold of less than 95% (Arlettaz 2006; Bhola 2014; Gomez-Rodriguez 2015; Meberg 2008; Oakley 2015; Richmond 2002; Turska 2012; Zuppa 2015), three studies used a post-ductal saturation threshold $\leq$ 95% (Koppel 2003; Riede 2010; Rosati 2005), and six studies used both pre-ductal and post-ductal saturations less than 95% (de-Wahl Granelli 2009; Ewer 2011; Klausner 2017; Singh 2014; Van Niekerk 2016; Zhao 2014). Two studies reported different positive thresholds (Bakr 2005 reported both pre-ductal and post-ductal fractional saturation $\leq$ 94%, and Sendelbach 2008 reported post-ductal saturation < 96%) (Table 1). In summary, the most common threshold was less than 95% (n = 14); five studies reported a threshold lower than or equal to 95%, and two studies reported thresholds $\leq$ 94% and < 96%, respectively. No study reported results for more than one threshold. Studies used different methods to verify test results: Investigators verified positive test results by echocardiography and negative results by interrogation of congenital anomaly registers, mortality data, or clinical follow-up (Table 1).

Methodological quality of included studies

We appraised the quality of primary diagnostic accuracy studies using the QUADAS-2 tool. We present quality assessment results for individual studies in the Characteristics of included studies table and in Figure 2. We have summarized the overall risk of bias and applicability concerns of studies in Figure 3.

Figure 2. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

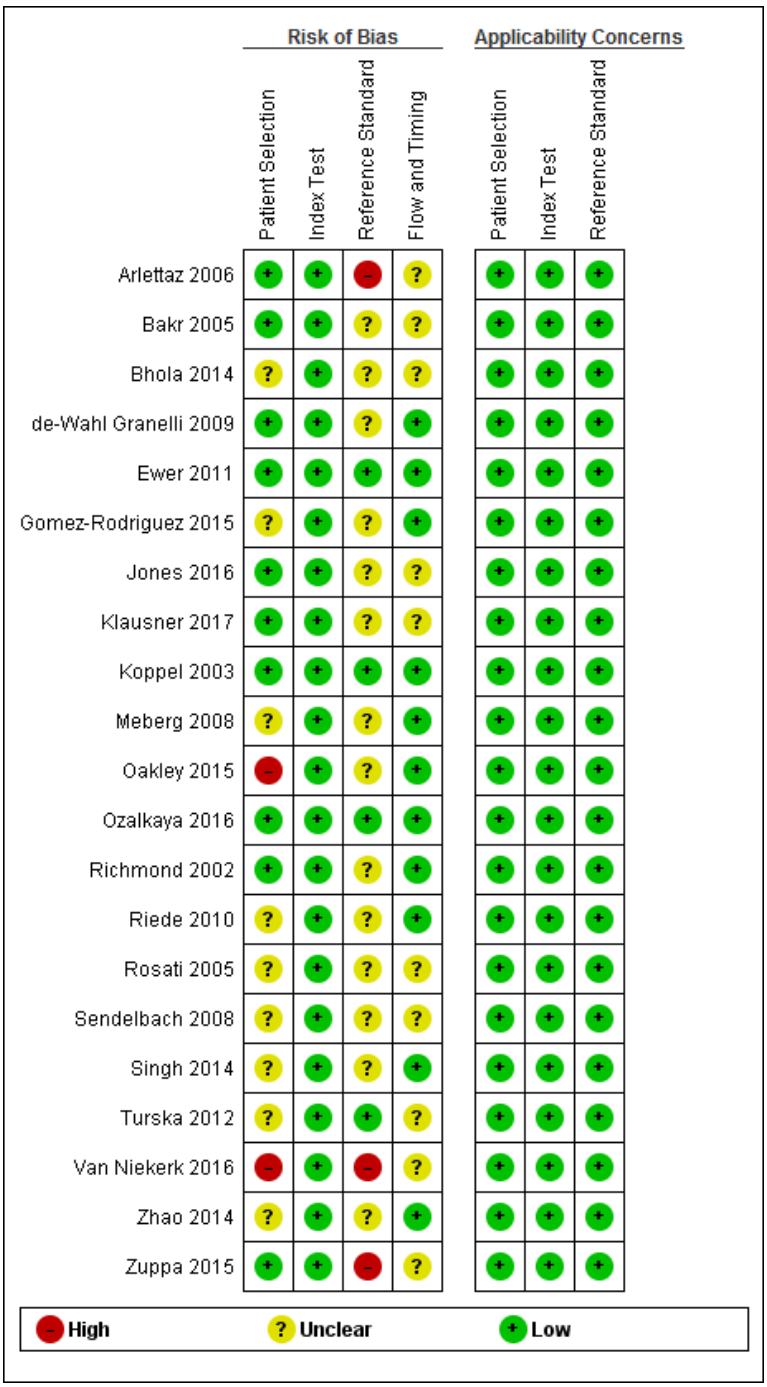
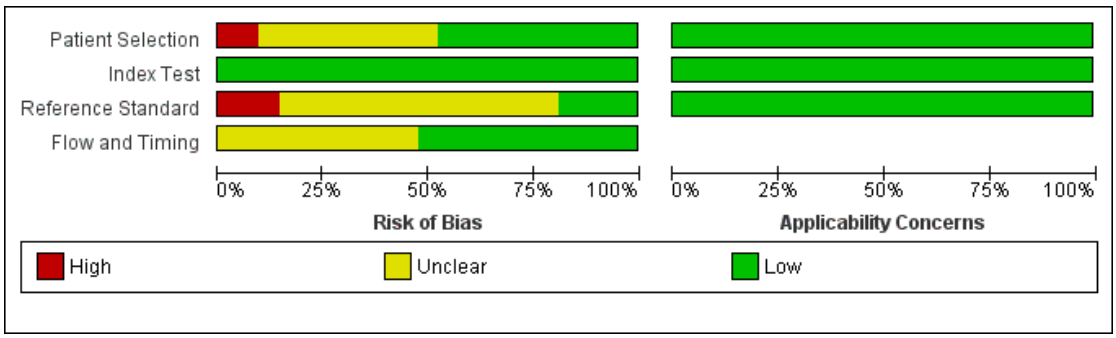


Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.



We judged the risk that patient selection (QUADAS-2, domain 1) had introduced bias as low in 10 studies (Arlettaz 2006; Bakr 2005; de-Wahl Granelli 2009; Ewer 2011; Jones 2016; Klausner 2017; Koppel 2003; Ozalkaya 2016; Richmond 2002; Zuppa 2015), high in two because investigators did not avoid inappropriate exclusions (Oakley 2015; Van Niekerk 2016), and unclear in the remaining nine studies (Bhola 2014; Gomez-Rodriguez 2015; Meberg 2008; Riede 2010; Rosati 2005; Sendelbach 2008; Singh 2014; Turska 2012; Zhao 2014). Applicability was of low concern for all studies in the patient selection domain.

For the index test assessment (QUADAS-2, domain 2), we considered all studies to be at low risk of bias and low concern regarding applicability.

We judged the risk that conduct or interpretation of reference standard(s) (QUADAS-2, domain 3) had introduced bias as low in four studies because investigators used echocardiography to confirm both positive and negative pulse oximetry cases (Ozalkaya 2016), or because they used echocardiography to confirm pulse oximetry positives and clinical follow-up in the first 28 days of life, which included postmortem findings and mortality and congenital anomaly databases to identify false-negative screening cases (Ewer 2011; Koppel 2003; Turska 2012). This comprehensive combination of clinical follow-up and review of registries and databases was considered as having low risk of bias. We considered that three studies reporting only echocardiography as the reference standard for positive pulse oximetry results were at high risk of bias (Arlettaz 2006; Van Niekerk 2016; Zuppa 2015). We considered risk for the remaining 14 studies as unclear because they used an incomplete reference standard to identify false-negative cases (Bakr 2005; Bhola 2014; de-Wahl Granelli 2009; Gomez-Rodriguez 2015; Jones 2016; Klausner 2017; Meberg 2008; Oakley 2015; Richmond 2002; Riede 2010; Rosati 2005; Sendelbach 2008; Singh 2014; Zhao 2014); six studies used echocardiography and follow-up (Gomez-Rodriguez 2015; Klausner 2017; Meberg 2009; Rosati 2005; Sendelbach 2008; Zhao 2014), and eight studies used echocardiography and different mortality and malformations registries (Bakr 2005; Bhola 2014; de-Wahl Granelli 2009; Jones 2016; Oakley 2015; Richmond 2002; Riede 2010; Singh 2014). It is noteworthy that only one study used echocardiography for positive and negative pulse oximetry results (Ozalkaya 2016). Applicability was of low concern for all studies in the reference standard(s) domain.

For flow and timing assessment (QUADAS-2, domain 4), 11 studies were at low risk of bias (de-Wahl Granelli 2009; Ewer 2011; Gomez-Rodriguez 2015; Koppel 2003; Meberg 2008; Oakley 2015; Ozalkaya 2016; Richmond 2002; Riede 2010; Singh 2014; Zhao 2014), and the remaining studies were at unclear risk because information reported was insufficient to permit judgment (Arlettaz 2006; Bakr 2005; Bhola 2014; Jones 2016; Klausner 2017; Rosati 2005; Sendelbach 2008; Turska 2012; Van Niekerk 2016; Zuppa 2015).

Findings

Results of meta-analysis

We considered for primary analysis all studies with thresholds around 95% (< 95% and ≤ 95%). As expected, this was the most common threshold among included studies (n = 19 studies; 436,758 participants) (Arlettaz 2006; Bhola 2014; de-Wahl Granelli 2009; Ewer 2011; Gomez-Rodriguez 2015; Jones 2016; Klausner 2017; Koppel 2003; Meberg 2008; Oakley 2015; Ozalkaya 2016; Richmond 2002; Riede 2010; Rosati 2005; Singh

2014; Turska 2012; Van Niekerk 2016; Zhao 2014; Zuppa 2015). The overall sensitivity of pulse oximetry for detection of critical congenital heart defects was 76.3% (95% confidence interval [CI] 69.5 to 82.0). Specificity was 99.9% (95% CI 99.7 to 99.9) with a false-positive rate of 0.14% (95% CI 0.07 to 0.22) (Summary of findings). Summary positive and negative likelihood ratios were 535.6 (95% CI 280.3 to 1023.4) and 0.24 (95% CI 0.18 to 0.31), respectively.

Fourteen out of 19 studies used a threshold lower than 95% (Arlettaz 2006; Bhola 2014; de-Wahl Granelli 2009; Ewer 2011; Gomez-Rodriguez 2015; Klausner 2017; Meberg 2008; Oakley 2015; Richmond 2002; Singh 2014; Turska 2012; Van Niekerk 2016; Zhao 2014; Zuppa 2015), and five studies used a threshold lower than or equal to 95% (Jones 2016; Koppel 2003; Ozalkaya 2016; Riede 2010; Rosati 2005).

Two additional studies used different thresholds: One used a threshold lower than or equal to 94% with sensitivity and specificity of 100% (95% CI 29 to 100) and 100% (95% CI 100 to 100), respectively (Bakr 2005); the other used a threshold of less than 96% with sensitivity and specificity of 100% (95% CI 3 to 100) and 100% (95% CI 100 to 100), respectively (Sendelbach 2008).

Overall, we have included in this review 349 cases of CCHD. The median prevalence of CCHD was 0.6 per 1000 live births (range 0.1 to 3.7; interquartile range 0.4 to 1.2).

Investigations of heterogeneity

To visualize total variability in sensitivity and specificity, we present the data in forest and ROC scatter plots (Figure 4; Figure 5). Forest plots show studies in increasing order of specificity (Figure 4). Sensitivity of the 21 studies ranged from 0% to 100%, and specificity from 99% to 100%. Forest and ROC plots show greater variability in estimated sensitivity than specificity across studies. Given results from investigations of heterogeneity, we used the bivariate model to estimate summary sensitivity and specificity (summary points) instead of the hierarchical summary ROC model to estimate summary ROC curves.

Figure 4. Forest plot of sensitivity and specificity. The figure shows the estimated sensitivity and specificity of the study (blue square) and its 95% confidence interval (black horizontal line). Studies are ordered by ascending specificity.

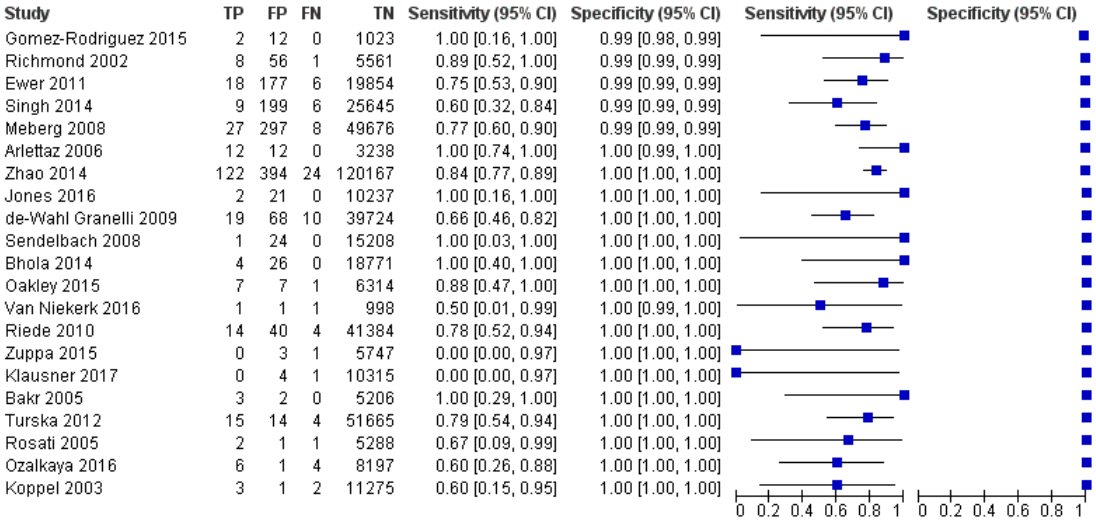
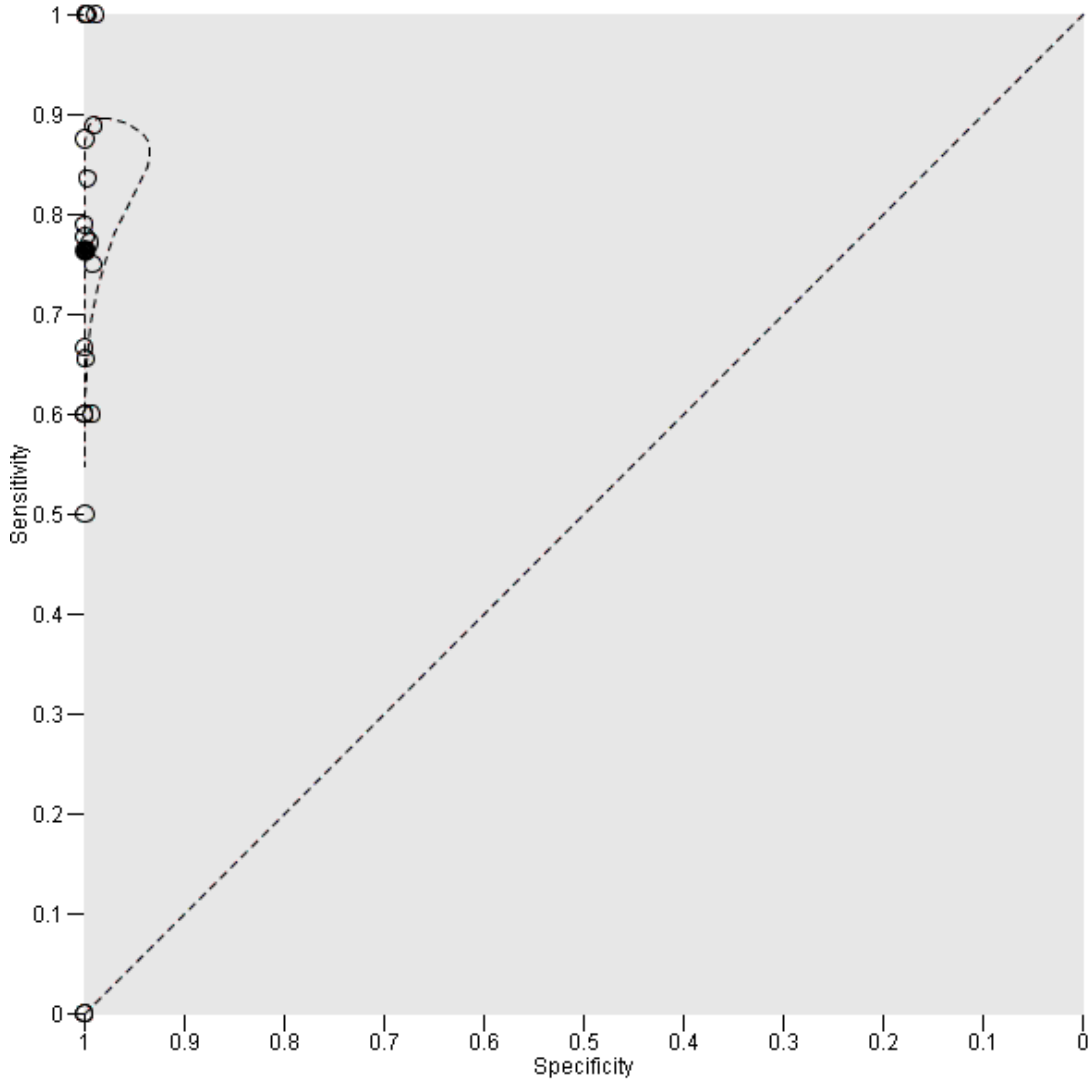


Figure 5. Summary ROC plot for pulse oximetry using a threshold lower than or lower than or equal to 95% (n = 19 studies). The solid circle corresponds to the summary estimate of sensitivity and specificity, and is shown with a 95% prediction region (dashed line).



For the primary analysis, we measured total between-study variability in sensitivity and in specificity through variances of the random effects for logit(sensitivity), logit(specificity), and their covariance, which were 0.102, 2.001, and -0.340, respectively. We represented the summary operating point with a 95% prediction region (Figure 5) and explored heterogeneity by differentiating studies on the basis of antenatal screening for CHD, timing of testing, type of oxygen saturation, study design, and risk of bias for the "flow and timing" domain of QUADAS-2. We plotted subgroups of studies in the ROC space.

Subgroup analysis and meta-regression

Table 2 summarizes results of the subgroup analysis including sensitivity and false-positive rates.

Antenatal diagnosis

Four studies included newborn infants with antenatal screening (Arlettaz 2006; Bhola 2014; Ewer 2011; Richmond 2002), and 15 studies did not (de-Wahl Granelli 2009; Gomez-Rodriguez 2015; Jones 2016; Klausner 2017; Koppel 2003; Meberg 2008; Oakley 2015; Ozalkaya 2016; Riede 2010; Rosati 2005; Singh 2014; Turska 2012; Van Niekerk 2016; Zhao 2014; Zuppa 2015). Summary estimates of sensitivity were 86.3% (95% CI 71.8 to 94.0) for studies that included antenatal screening, and 74.1% (95% CI 65.7 to 81.1) for studies that did not include antenatal screening. Summary estimates of specificity were 99.5% (95% CI 98.4 to 99.9) with a false-positive rate of 0.46% (95% CI 0.13 to 1.59) for studies with antenatal screening, and 99.9% (95% CI 99.8 to 100) with a false positive rate of 0.10% (95% CI 0.05 to 0.21) for studies that did not include antenatal screening. Sensitivity (P = 0.071) and specificity (P = 0.231) did not change significantly when newborn infants with antenatal suspicion of congenital heart defects were included compared with when they were excluded.

Test timing

Eleven studies performed pulse oximetry screening after 24 hours from birth (Bhola 2014; de-Wahl Granelli 2009; Klausner 2017; Koppel 2003; Oakley 2015; Ozalkaya 2016; Riede 2010; Rosati 2005; Van Niekerk 2016; Zhao 2014; Zuppa 2015), and the other eight studies performed pulse oximetry within 24 hours of birth (Arlettaz 2006; Ewer 2011; Gomez-Rodriguez 2015; Jones 2016; Meberg 2008; Richmond 2002; Singh 2014; Turska 2012). Summary estimates of sensitivity and specificity of studies that performed screening after 24 hours were 73.6% (95% CI 62.8 to 82.1) and 99.9% (95% CI 99.9 to 100). For studies that performed screening within 24 hours, summary estimates of sensitivity and specificity were 79.5% (95% CI 70.0 to 86.6) and 99.6% (95% CI 99.1 to 99.8). Test timing to perform pulse oximetry had no significant effect on sensitivity (P = 0.393), but the false-positive rate for detection of CCHD was lower when newborn pulse oximetry was done after 24 hours from birth than when it was done within 24 hours (0.06% [95% CI 0.03 to 0.13] vs 0.42% [95% CI 0.20 to 0.89]; P = 0.027).

Limbs

Eleven studies used the foot alone (post-ductal) to measure oxygen saturation (Arlettaz 2006; Bhola 2014; Gomez-Rodriguez 2015; Koppel 2003; Meberg 2008; Oakley 2015; Richmond 2002; Riede 2010; Rosati 2005; Turska 2012; Zuppa 2015); summary estimates of sensitivity and specificity were 81.2% (95% CI 70.9 to 88.4) and 99.9% (95% CI 99.7 to 100), respectively, with a false-positive rate of 0.13% (95% CI 0.05 to 0.31). Eight studies used both right hand and foot (pre-ductal and post-ductal) (de-Wahl Granelli 2009; Ewer 2011; Jones 2016; Klausner 2017; Ozalkaya 2016; Singh 2014; Van Niekerk 2016; Zhao 2014); summary estimates of sensitivity and specificity for this group of studies were 71.2% (95% CI 58.5 to 81.3) and 99.8% (95% CI 99.5 to 99.9), respectively, with a false-positive rate of 0.17% (95% CI 0.06 to 0.46). We noted no significant differences in sensitivity (P = 0.197) nor in specificity (P = 0.718) for pulse oximetry when measures were obtained in the foot alone rather than in both the foot and the right hand.

Risk of bias

We judged nine studies as having unclear risk of bias for the "flow and timing" domain of QUADAS-2 (Arlettaz 2006; Bhola 2014; Klausner 2017; Koppel 2003; Riede 2010; Rosati 2005; Turska 2012; Van Niekerk 2016; Zuppa 2015). Summary estimates of sensitivity and specificity were 77.8% (95% CI 64.1 to 87.3) and 100% (95% CI 99.9 to 100), respectively, with a false-positive rate of 0.05% (95% CI 0.02 to 0.12). We judged the remaining 10 studies as having low risk of bias (de-Wahl Granelli 2009; Ewer 2011; Gomez-Rodriguez 2015; Jones 2016; Meberg 2008; Oakley 2015; Ozalkaya 2016; Richmond 2002; Singh 2014; Zhao 2014); summary estimates of sensitivity and specificity were 77.3% (95% CI 68.8 to 84.0) and 99.7% (95% CI 99.3 to 99.8), respectively, with a false-positive rate of 0.34% (95% CI 0.17 to 0.66). Risk of bias for this domain had no significant effect on sensitivity (P = 0.937), but studies judged as having unclear risk of bias for the "flow and timing" domain had higher specificity (P = 0.016).

Sensitivity analysis

We performed a sensitivity analysis while excluding from the primary analysis studies with a threshold $\leq$ 95% (Jones 2016; Koppel

2003; Ozalkaya 2016; Riede 2010; Rosati 2005). For this analysis, sensitivity and specificity were 78.1% (95% CI 71.2 to 83.7) and 99.8% (95% CI 99.6 to 99.9) with a false-positive rate of 0.23% (95% CI 0.12 to 0.44). Exclusion of these studies increased the sensitivity and false-positive rate of pulse oximetry screening.

We also performed a sensitivity analysis for which we added to the primary analysis studies with a threshold ≤ 94% and < 96% (Bakr 2005; Sendelbach 2008). For this analysis, sensitivity and specificity were 77% (95% CI 70 to 82) and 100% (95% CI 100 to 100), respectively. Inclusion of these studies produced a slight improvement in the sensitivity of the test.

Furthermore, we investigated the effects of potential sources of bias by removing the four studies judged as having high risk of bias in one of the QUADAS-2 domains (Arlettaz 2006; Oakley 2015; Van Niekerk 2016; Zuppa 2015). For this analysis, sensitivity and specificity were similar to those in the main analysis, at 75.5% (95% CI 68.2 to 81.6) and 99.79% (95% CI 99.7 to 99.9), respectively.

Summary of findings

Should pulse oximetry be used to diagnose CCHD in asymptomatic newborns?

Patient or population: asymptomatic newborns at the time of pulse oximetry screening

Setting: hospital births

Index test: pulse oximetry

Reference test: Reference standards were both diagnostic echocardiography (echocardiogram) and clinical follow-up in the first 28 days of life, including postmortem findings and mortality and congenital anomaly databases to identify false-negative patients

Studies: We included prospective or retrospective cohorts and cross-sectional studies. We excluded case reports and studies of case-control design

Threshold	Summary accuracy (95% CI)	Number of participants (diseased/non-diseased) Number of studies	Prevalence median (range)	Implications (in a cohort of 10,000 newborns tested [95% CI])				Certainty of the evidence (GRADE)
					Prevalence 0.6 per 1000	Prevalence 0.1 per 1000	Prevalence 3.7 per 1000	
95% (less than or less than or equal to)	**Sensitivity** 76.3% (69.5 to 82.0) **Specificity** 99.9% (99.7 to 99.9)	436,758 (345/436,413) 19 studies	0.6 per 1000 (0.1 to 3.7)	**True positives** *(newborns with CCHD)*	5 (4 to 5)	1 (1 to 1)	28 (26 to 30)	**LOW*** ⊕⊕◯◯
				False negatives *(newborns incorrectly classified as not having CCHD)*	1 (1 to 2)	0 (0 to 0)	9 (7 to 11)	
				True negatives *(newborns without CCHD)*	9980 (9966 to 9987)	9985 (9971 to 9992)	9949 (9935 to 9956)	**HIGH** ⊕⊕⊕⊕

	False positives *(newborns incorrectly classified as having CCHD)*	14 (7 to 28)	14 (7 to 28)	14 (7 to 28)

CCHD: critical congenital heart defect; CI: confidence interval

Sensitivity:

**We have downgraded certainty of the evidence from high to low because the low number of CCHD cases included in the review (serious imprecision) and secondly, there was a serious risk of differential verification bias (ie, diagnosis was established by echocardiography in test positive cases however test negatives were usually confirmed by clinical follow-up or by accessing congenital malformation registries and mortality databases)."*

Certainty of the evidence (Balshem 2011)

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

DISCUSSION

Summary of main results

For this review, we have identified and summarized the results of all available cohort studies reporting the test accuracy of pulse oximetry screening for detection of critical congenital heart defects (CCHDs) in asymptomatic late preterm and full-term infants in postnatal wards or well-baby nurseries. We have presented the main results in Summary of findings. We analyzed data on 457,202 participants from 21 included studies. We restricted the primary analysis to studies with thresholds around 95% (< 95% and ≤ 95%). Analysis, including 436,758 participants from 19 studies, showed that pulse oximetry screening is a highly specific screening test with moderate sensitivity and a low overall false-positive rate. Overall sensitivity was 76.3%, specificity was 99.9%, and the false-positive rate was 0.14%. Summary positive and negative likelihood ratios were 535.6 and 0.24, respectively. Inclusion of studies that used different saturation thresholds from those in the primary analysis slightly improved the sensitivity of the test. Exclusion of studies at high risk of bias did not significantly alter overall sensitivity or specificity. Between-study heterogeneity was higher in sensitivity than in specificity estimates.

Most studies were conducted in high-income countries (USA, Europe); however, we also included studies from middle-income countries, which increases the generalizability of review findings. We noted methodological variation between studies with respect to inclusion or exclusion of babies with a suspected antenatal diagnosis, timing of testing (before or after 24 hours of age), site of testing (post-ductal only or pre-ductal and post-ductal), functional or fractional saturation measurement, and study design (prospective or retrospective). Subgroup analysis showed no effect on sensitivity or specificity among these variables, although later screening was associated with a lower false-positive rate than was reported with earlier screening.

The definition of CCHD provided in the published literature is highly variable. We attempted to address this by applying a strict definition (see above) to categorize CCHD in a standardized manner, thus reducing the risk of an incorrect diagnosis.

Strengths and weaknesses of the review

Strengths of this review include a comprehensive literature search performed to identify all relevant studies, rigorous assessment of risk of bias of included studies using the QUADAS-2 tool, duplicate data extraction, and performance of subgroup and sensitivity analyses to investigate differences in estimates of accuracy of pulse oximetry among studies with high, low, or unclear risk of bias. However, only one study included more than 100 CCHD cases, and 12 studies included fewer than 10 cases. The relatively low number of CCHD cases included in this review indicates that the precision of sensitivity is still low.

Our review has explored and quantified the heterogeneity, and review authors have tried to identify possible sources of heterogeneity. Exploration of sources of heterogeneity has produced different results for sensitivity and specificity. Sensitivity has not been affected by any of the a priori selected sources of heterogeneity. We cannot rule out the presence of unexplained heterogeneity in this accuracy index, although it is highly likely that some of the variability observed in sensitivities of individual studies could be explained by the paucity of CCHD cases. Use of different strategies for confirming pulse oximetry negative cases (ie, passively with mortality or registry data rather than active clinical follow-up) could well have introduced some degree of heterogeneity into sensitivity results. However, this post hoc exploration was not performed, given the scarcity of data. This means that sensitivity estimates are somewhat unstable with wide confidence intervals. At the same time, this scarcity made analysis of heterogeneity underpowered. Conversely, specificity was affected by the timing of the test and by the risk of bias due to the flow and timing domain of the QUADAS-2 tool. Statistical significance achieved by the specificity analysis is a direct consequence of the large number of healthy newborns included in the review. On the other hand, the magnitude of differences between subgroup analyses was small. False-positive rates were 0.06% and 0.42% for newborns screened after and before 24 hours of birth, respectively. The absolute difference was 0.36% with more false-positives in the earlier screening group as compared with the late screening group. This means, in relative terms, seven times more false positives are seen in the earlier screening group than in the late screening group. Similarly, false-positive rates varied between studies judged as having unclear or low risk of bias for the "flow and timing" domain of QUADAS-2 (ie, 0.05% vs 0.34% for unclear and low risk groups of studies, respectively). The absolute difference a 0.29% reduction in false positives in the unclear risk group, which equates almost seven times fewer false positives in relative terms.

Agreements and disagreements with other studies or reviews

This review includes more studies and a larger body of data from a significantly greater number of infants than were included in similar previous systematic reviews of the test accuracy of pulse oximetry screening to detect CCHD (Mahle 2009; Thangaratinam 2007; Thangaratinam 2012), which reported identical statistical methods and meta-analyses. Compared with the largest prior review (Thangaratinam 2012), authors of this review screened a significantly larger number of references (2695 vs 552) and included data from over 220,000 more babies, allowing greater precision of the estimates of test accuracy, and providing the most complete meta-analysis available so far.

Overall sensitivity is similar (76.3% vs 76.5%) to that described by Thangaratinam 2012 and is similar to the overall false-positive

rate (0.14% vs 0.14%). The statistically significant lower false-positive rate between early and later screening persists (0.06% vs 0.42% and 0.05% vs 0.5%).

Applicability of findings to the review question

This review includes a large number of relevant studies that met our inclusion criteria, and review authors had few concerns about the relevance of their findings to our review questions. We mainly judged included studies to be at low or unclear risk of bias in QUADAS-2 domains. Most studies had a prospective design with consecutive enrollment of participants and included an adequate description of the index test. Some studies reported the exclusion criteria poorly. Data were complete and were available for all included studies.

Risk of differential verification bias was unavoidable as diagnosis was established by echocardiography in test-positive cases; however, test-negative cases were usually confirmed by clinical follow-up or by examination of congenital malformation registries and mortality databases; risk of bias in the conduct or interpretation of reference standard(s) was unclear in most studies that used incomplete reference standards. This of course raises the possibility that some of the false negatives may be misclassified as true negatives. This misclassification overestimates sensitivity and specificity. Differential verification may have had an impact on the sensitivity estimate. For this reason, along with the potential for imprecision, given the small number of CCHD cases, we have downgraded the GRADE certainty of evidence for sensitivity to "low." In our review, studies judged as having unclear risk of bias for the "flow and timing" domain showed higher specificity.

AUTHORS' CONCLUSIONS

Implications for practice

This review provides further compelling evidence for the use of pulse oximetry as a routine screening test for early identification of CCHD in asymptomatic babies in the well-baby nursery. The test has high specificity and moderate sensitivity and meets the criteria for universal screening.

Current evidence supports the introduction of routine screening for CCHD in asymptomatic newborns before discharge from the well-baby nursery. The test appears feasible in various middle-income countries and shows consistent test accuracy.

Some important elements regarding specific screening algorithms need further consideration. Data show no difference in sensitivity based on the site of testing (pre-ductal or pre-ductal and post-ductal). However, only two studies using pre-ductal and post-ductal saturations reported absolute saturation values rather than just test results (de-Wahl Granelli 2009; Ewer 2011). As has been reported previously by Ewer 2016 and Thangaratinam 2012, several CCHD cases that were detected by pre-ductal and post-ductal testing would have been missed by post-ductal testing alone, but the numbers are too small to affect sensitivity analysis results.

In addition, the finding of a lower false-positive rate with screening after 24 hours needs to be balanced against the fact that many countries discharge babies within 24 hours and - as is important to note - most reported studies did not take into account the risk that a baby with CCHD or other serious illness may present before screening takes place (de-Wahl Granelli 2014; Ewer 2011; Ewer 2016; Riede 2010; Thangaratinam 2012).

The prevalence of CCHD is quite low, and most test-positive infants do not have the target condition. The false-positive rate is variable and depends largely on the timing of the screening (earlier screening - within 24 hours of age - has a higher false-positive rate than screening after 24 hours). This raises concerns that a false-positive test may unnecessarily increase parental anxiety and may lead to avoidable investigations and delayed discharge. Investigators in the UK PulseOx study assessed the acceptability of pulse oximetry screening and reported on anxiety created by the test - particularly among mothers of false-positive (FP) babies (Ewer 2012a; Powell 2013). Investigators quantified satisfaction with, and perceptions of, the test and anxiety and depression following screening by using validated questionnaires on samples of mothers whose babies were true positive, false positive, and true negative. All participants were predominantly satisfied with screening, and it is important to note that mothers given false-positive results after screening were no more anxious than those given true-negative results. Many studies report identification of alternative non-cardiac conditions by pulse oximetry screening. Although these conditions - such as congenital pneumonia and early-onset sepsis - are technically false positives, their identification may be seen as a positive additional benefit of screening; they are more likely to be detected within the first 24 hours, allowing early treatment of individuals with these potentially serious conditions. Healthcare providers must consider the potential for overdiagnosis of these conditions following screening and must apply rigorous criteria to classify these conditions.

Implications for research

The large sample size of this review along with precise estimates of sensitivity and specificity suggests that further research into the accuracy of this screening method is unnecessary. In addition, several countries, including the USA, have already implemented screening. However, given concerns related to differential verification, we propose that monitoring of screening outcomes (including possible reduction in early mortality) and management of false positives should be performed in a rigorous manner.

Further evidence regarding the routine screening of babies outside

the well-baby nursery (including non-intensive care unit [NICU] stays and out-of-hospital births) is required. Additional raw saturation data and further analysis are required to further elucidate the relative sensitivities of post-ductal versus pre-ductal and post-ductal saturation testing.

The ability of pulse oximetry to detect non-cardiac illness such as respiratory and infectious conditions has been well described, but test accuracy remains unclear.

ACKNOWLEDGEMENTS

We would like to thank Colleen Ovelman (Managing Editor and Assistant Information Specialist), Yolanda Brosseau (Webmaster), and Jennifer Spano (Information Specialist) of the Neonatal Group for their kind and efficient support during preparation of this review.

REFERENCES

References to studies included in this review

Arlettaz 2006 *{published data only}*
Arlettaz R, Bauschatz AS, Monkhoff M, Essers B, Bauersfeld U. The contribution of pulse oximetry to the early detection of congenital heart disease in newborns. *European Journal of Pediatrics* 2006;**165**(2):94–8. PUBMED: 10.1007/s00431–005–0006–y; PUBMED: 16211399]

Bakr 2005 *{published data only}*
Bakr AF, Habib HS. Combining pulse oximetry and clinical examination in screening for congenital heart disease. *Pediatric Cardiology* 2005;**26**(6):832–5. [DOI: 10.1007/s00246-005-0981-9; PUBMED: 16088415

Bhola 2014 *{published data only}*
Bhola K, Kluckow M, Evans N. Post-implementation review of pulse oximetry screening of well newborns in an Australian tertiary maternity hospital. *Journal of Paediatrics and Child Health* 2014;**50**(11):920–5. [DOI: 10.1111/jpc.12651; PUBMED: 24923996

de-Wahl Granelli 2009 *{published data only}*
de-Wahl Granelli A, Wennergren M, Sandberg K, Mellander M, Bejlum C, Inganas L, et al. Impact of pulse oximetry screening on the detection of duct dependent congenital heart disease: a Swedish prospective screening study in 39, 821 newborns. *BMJ (Clinical Research Ed.)* 2009;**338**:a3037. [DOI: 10.1136/bmj.a3037; PUBMED: 19131383

Ewer 2011 *{published data only}*
Ewer AK, Middleton LJ, Furmston AT, Bhoyar A, Daniels JP, Thangaratinam S, et al. PulseOx Study Group. Pulse oximetry screening for congenital heart defects in newborn infants (PulseOx): a test accuracy study. *Lancet* 2011;**378**(9793):785–94. [DOI: 10.1016/S0140-6736(11)60753-8; PUBMED: 21820732

Gomez-Rodriguez 2015 *{published data only}*
Gomez-Rodriguez G, Quezada-Herrera A, Amador-Licona N, Carballo-Magdaleno D, Rodriguez-Mejia EJ, Guizar-Mendoza JM. Pulse oximetry as a screening test for critical congenital heart disease in term newborns. *Revista de Investigacion Clinica* 2015;**67**(2):130–4. PUBMED: 25938847]

Jones 2016 *{published data only}*
Jones AJ, Howarth C, Nicholl R, Mat-Ali E, Knowles R. The impact and efficacy of routine pulse oximetry screening for CHD in a local hospital. *Cardiology in the Young* 2016;**26**(7):1397–405. [DOI: 10.1017/S1047951115002784; PUBMED: 26905447

Klausner 2017 *{published data only}*
Klausner R, Shapiro ED, Elder RW, Colson E, Loyal J. Evaluation of a screening program to detect critical congenital heart defects in newborns. *Hospital Pediatrics* 2017;**7**(4):214–8. [DOI: 10.1542/hpeds.2016-0176; PUBMED: 28250095

Koppel 2003 *{published data only}*
Koppel RI, Druschel CM, Carter T, Goldberg BE, Mehta PN, Talwar R, et al. Effectiveness of pulse oximetry screening for congenital heart disease in asymptomatic newborns. *Pediatrics* 2003;**111**(3):451–5. PUBMED: 12612220]

Meberg 2008 *{published data only}*
Meberg A, Brugmann-Pieper S, Due R Jr, Eskedal L, Fagerli I, Farstad T, et al. First day of life pulse oximetry screening to detect congenital heart defects. *Journal of Pediatrics* 2008;**152**(6):761–5. [DOI: 10.1016/j.jpeds.2007.12.043; PUBMED: 18492511

Oakley 2015 *{published data only}*
Oakley JL, Soni NB, Wilson D, Sen S. Effectiveness of pulse-oximetry in addition to routine neonatal examination in detection of congenital heart disease in asymptomatic newborns. *Journal of Maternal-Fetal & Neonatal Medicine* 2015;**28**(14):1736–9. [DOI: 10.3109/14767058.2014.967674; PUBMED: 25241768

Ozalkaya 2016 *{published data only}*
Ozalkaya E, Akdag A, Sen I, Comert E, Melek Yaren H. Early screening for critical congenital heart defects in asymptomatic newborns in Bursa province. *Journal of Maternal-Fetal & Neonatal Medicine* 2016;**29**(7):1105–7. [DOI: 10.3109/14767058.2015.1035642; PUBMED: 25902399

Richmond 2002 *{published data only}*
Richmond S, Reay G, Abu Harb M. Routine pulse oximetry in the asymptomatic newborn. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2002;**87**(2):F83–8. PUBMED: 12193511]

Riede 2010 *{published data only}*
Riede FT, Worner C, Dahnert I, Mockel A, Kostelka M, Schneider P. Effectiveness of neonatal pulse oximetry

screening for detection of critical congenital heart disease in daily clinical routine - results from a prospective multicenter study. *European Journal of Pediatrics* 2010; **169**(8):975–81. [DOI: 10.1007/s00431-010-1160-4; PUBMED: 20195633

Rosati 2005 *{published data only}*
Rosati E, Chitano G, Dipaola L, De Felice C, Latini G. Indications and limitations for a neonatal pulse oximetry screening of critical congenital heart disease. *Journal of Perinatal Medicine* 2005;**33**(5):455–7. [DOI: 10.1515/JPM.2005.080; PUBMED: 16238542

Sendelbach 2008 *{published data only}*
Sendelbach DM, Jackson GL, Lai SS, Fixler DE, Stehel EK, Engle WD. Pulse oximetry screening at 4 hours of age to detect critical congenital heart defects. *Pediatrics* 2008;**122**(4):e815–20. [DOI: 10.1542/peds.2008-0781; PUBMED: 18762486

Singh 2014 *{published data only}*
Singh A, Rasiah SV, Ewer AK. The impact of routine predischarge pulse oximetry screening in a regional neonatal unit. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2014;**99**(4):F297–302. [DOI: 10.1136/archdischild-2013-305657; PUBMED: 24646619

Turska 2012 *{published data only}*
Turska Kmiec A, Borszewska Kornacka MK, Blaz W, Kawalec W, Zuk M. Early screening for critical congenital heart defects in asymptomatic newborns in Mazovia province: experience of the POLKARD pulse oximetry programme 2006-2008 in Poland. *Kardiologia Polska* 2012;**70**(4):370–6. 22528711]

Van Niekerk 2016 *{published data only}*
Van Niekerk AM, Cullis RM, Linley LL, Zuhlke L. Feasibility of pulse oximetry pre-discharge screening implementation for detecting critical congenital heart lesions in newborns in a secondary level maternity hospital in the Western Cape, South Africa: The 'POPSICLe' study. *South African Medical Journal* 2016;**106**(8):817–21. PUBMED: 27499412]

Zhao 2014 *{published data only}*
Zhao QM, Ma XJ, Ge XL, Liu F, Yan WL, Wu L, et al. Neonatal Congenital Heart Disease Screening Group. Pulse oximetry with clinical assessment to screen for congenital heart disease in neonates in China: a prospective study. *Lancet* 2014;**384**(9945):747–54. [DOI: 10.1016/S0140-6736(14)60198-7; PUBMED: 24768155

Zuppa 2015 *{published data only}*
Zuppa AA, Riccardi R, Catenazzi P, D'Andrea V, Cavani M, D'Antuono A, et al. Clinical examination and pulse oximetry as screening for congenital heart disease in low-risk newborn. *Journal of Maternal-Fetal & Neonatal Medicine* 2015;**28**(1):7–11. [DOI: 10.3109/14767058.2014.899573; PUBMED: 24588079

References to studies excluded from this review

Andrews 2014 *{published data only}*
Andrews JP, Ross AS, Salazar MA, Tracy NA, Burke BL Jr. Smooth implementation of critical congenital heart defect screening in a newborn nursery. *Clinical Pediatrics* 2014;**53**(2):173–6. [DOI: 10.1177/0009922813502850; PUBMED: 24037922

Ewer 2012a *{published data only}*
Ewer AK, Furmston AT, Middleton LJ, Deeks JJ, Daniels JP, Pattison HM, et al. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a test accuracy study with evaluation of acceptability and cost-effectiveness. *Health Technology Assessment (Winchester, England)* 2012;**16**(2):v-xiii, 1-184. [DOI: 10.3310/hta16020; PUBMED: 22284744

Hoke 2002 *{published data only}*
Hoke TR, Donohue PK, Bawa PK, Mitchell RD, Pathak A, Rowe PC, et al. Oxygen saturation as a screening test for critical congenital heart disease: a preliminary study. *Pediatric Cardiology* 2002;**23**(4):403–9. [DOI: 10.1007/s00246-002-1482-8; PUBMED: 12170356

John 2016 *{published data only}*
John C, Phillips J, Hamilton C, Lastliger A. Implementing universal pulse oximetry screening in West Virginia: findings from year one. *West Virginia Medical Journal* 2016;**112**(4):42–6. PUBMED: 27491102]

Kardasevic 2016 *{published data only}*
Kardasevic M, Jovanovic I, Samardzic JP. Modern strategy for identification of congenital heart defects in the neonatal period. *Medical Archives* 2016;**70**(5):384–8. [DOI: 10.5455/medarh.2016.70.384-388; PUBMED: 27994302

Kochilas 2013 *{published data only}*
Kochilas LK, Lohr JL, Bruhn E, Borman-Shoap E, Gams BL, Pylipow M, et al. Implementation of critical congenital heart disease screening in Minnesota. *Pediatrics* 2013;**132**(3):e587–94. [DOI: 10.1542/peds.2013-0803; PUBMED: 23958775

Lhost 2014 *{published data only}*
Lhost JJ, Goetz EM, Belling JD, van Roojen WM, Spicer G, Hokanson JS. Pulse oximetry screening for critical congenital heart disease in planned out-of-hospital births. *Journal of Pediatrics* 2014;**165**(3):485–9. [DOI: 10.1016/j.jpeds.2014.05.011; PUBMED: 24948344

Meberg 2009 *{published data only}*
Meberg A, Andreassen A, Brunvand L, Markestad T, Moster D, Nietsch L, et al. Pulse oximetry screening as a complementary strategy to detect critical congenital heart defects. *Acta Paediatrica* 2009;**98**(4):682–6. [DOI: 10.1111/j.1651-2227.2008.01199.x; PUBMED: 19154526

Movahedian 2016 *{published data only}*
Movahedian AH, Mosayebi Z, Sagheb S. Evaluation of pulse oximetry in the early detection of cyanotic congenital heart disease in newborns. *Journal of Tehran Heart Center* 2016;**11**(2):73–8. PUBMED: 27928258]

Narayen 2016a *{published data only}*
Narayen IC, Blom NA, Bourgonje MS, Haak MC, Smit M, Posthumus F, et al. Pulse oximetry screening for critical congenital heart disease after home birth and early discharge. *Journal of Pediatrics* 2016;**170**:188–92.e1. PUBMED: 10.1016/j.jpeds.2015.12.004]

Prudhoe 2013 *{published data only}*
Prudhoe S, Abu-Harb M, Richmond S, Wren C. Neonatal screening for critical cardiovascular anomalies using pulse oximetry. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2013;**98**(4):F346–50. [DOI: 10.1136/archdischild-2012-302045; PUBMED: 23341250

Reich 2003 *{published data only}*
Reich JD, Miller S, Brogdon B, Casatelli J, Gompf TC, Huhta JC, et al. The use of pulse oximetry to detect congenital heart disease. *Journal of Pediatrics* 2003;**142**(3):268–72. [DOI: 10.1067/mpd.2003.87; PUBMED: 12640374

Reich 2008 *{published data only}*
Reich JD, Connolly B, Bradley G, Littman S, Koeppel W, Lewycky P, et al. The reliability of a single pulse oximetry reading as a screening test for congenital heart disease in otherwise asymptomatic newborn infants. *Pediatric Cardiology* 2008;**29**(5):885–9. [DOI: 10.1007/s00246-008-9214-3; PUBMED: 18347842

Reich 2008a *{published data only}*
Reich JD, Connolly B, Bradley G, Littman S, Koeppel W, Lewycky P, et al. Reliability of a single pulse oximetry reading as a screening test for congenital heart disease in otherwise asymptomatic newborn infants: the importance of human factors. *Pediatric Cardiology* 2008;**29**(2):371–6. [DOI: 10.1007/s00246-007-9105-z; PUBMED: 17932712

Riede 2009 *{published data only}*
Riede FT, Dahnert I, Schneider P, Mockel A. Pulse oximetry screening at 4 hours of age to detect critical congenital heart defects. Pediatrics 2009; Vol. 123, issue 3:e542. [DOI: 10.1542/peds.2008-3598; PUBMED: 19254989

Ruangritnamchai 2007 *{published data only}*
Ruangritnamchai C, Bunjapamai W, Pongpanich B. Pulse oximetry screening for clinically unrecognized critical congenital heart disease in the newborns. *Images in Paediatric Cardiology* 2007;**9**(1):10–5. PUBMED: 22368668]

Saha 2014 *{published data only}*
Saha A, Mathew JL, Chawla D, Kumar D. How useful is pulse oximetry for screening of congenital heart disease in newborns?. *Indian Pediatrics* 2014;**51**(11):913–5. PUBMED: 25432224]

Saxena 2015 *{published data only}*
Saxena A, Mehta A, Ramakrishnan S, Sharma M, Salhan S, Kalaivani M, et al. Pulse oximetry as a screening tool for detecting major congenital heart defects in Indian newborns. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2015;**100**(5):F416–21. [DOI: 10.1136/archdischild-2014-307485; PUBMED: 26038347

Schena 2017 *{published data only}*
Schena F, Picciolli I, Agosti M, Zuppa AA, Zuccotti G, Parola L, et al. Neonatal Cardiology Study Group of the Italian Society of Neonatology. Perfusion index and pulse oximetry screening for congenital heart defects. *Journal of Pediatrics* 2017;**183**:74–9.e1. [DOI: 10.1016/j.jpeds.2016.12.076; PUBMED: 28153478

Studer 2014 *{published data only}*
Studer MA, Smith AE, Lustik MB, Carr MR. Newborn pulse oximetry screening to detect critical congenital heart disease. *Journal of Pediatrics* 2014;**164**(3):505–9.e1-2. [DOI: 10.1016/j.jpeds.2013.10.065; PUBMED: 24315501

Taksande 2013 *{published data only}*
Taksande AM, Lakhkar B, Gadekar A, Suwarnakar K, Japzape T. Accuracy of pulse oximetry screening for detecting critical congenital heart disease in the newborns in rural hospital of Central India. *Images in Paediatric Cardiology* 2013;**15**(4):5–10. PUBMED: 26236364]

Tautz 2010 *{published data only}*
Tautz J, Merkel C, Loersch F, Egen O, Hagele F, Thon HM, et al. Implication of pulse oxymetry screening for detection of congenital heart defects. *Klinische Padiatrie* 2010;**222**(5):291–5. [DOI: 10.1055/s-0030-1253391; PUBMED: 20458668

Tsao 2016 *{published data only}*
Tsao PC, Shiau YS, Chiang SH, Ho HC, Liu YL, Chung YF, et al. Development of a newborn screening program for critical congenital heart disease (CCHD) in Taipei. *PloS One* 2016;**11**(4):e0153407. [DOI: 10.1371/journal.pone.0153407; PUBMED: 27073996

Vaidyanathan 2011 *{published data only}*
Vaidyanathan B, Sathish G, Mohanan ST, Sundaram KR, Warrier KK, Kumar RK. Clinical screening for congenital heart disease at birth: a prospective study in a community hospital in Kerala. *Indian Pediatrics* 2011;**48**(1):25–30. PUBMED: 20972295]

Valmari 2006 *{published data only}*
Valmari P, Jauhola R, Leskinen M, Heikinheimo M. [Oxygen saturation to detect asymptomatic congenital heart disease of newborns] [Vastasyntyneen kavalat sydanviat—-aika aloittaa happikyllasteisyysseulonta?]. *Duodecim; Laaketieteellinen Aikakauskirja* 2006;**122**(2):173–6. PUBMED: 16509060]

Walsh 2011 *{published data only}*
Walsh W. Evaluation of pulse oximetry screening in Middle Tennessee: cases for consideration before universal screening. *Journal of Perinatology* 2011;**31**(2):125–9. [DOI: 10.1038/jp.2010.70; PUBMED: 20508595

Additional references

Abu-Harb 1994
Abu-Harb M, Hey E, Wren C. Death in infancy from unrecognised congenital heart disease. *Archives of Disease in Childhood* 1994;**71**(1):3–7. [PUBMED: 8067789]

Abu-Harb 1994a
Abu-Harb M, Wyllie J, Hey E, Richmond S, Wren C. Presentation of obstructive left heart malformations in infancy. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 1994;**71**(3):F179–83. [PUBMED: 7820713]

Allan 1986
Allan LD, Crawford DC, Chita SK, Tynan MJ. Prenatal screening for congenital heart disease. *British Medical Journal (Clinical Research Ed.)* 1986;**292**(6537):1717–9. [PUBMED: 3089369]

Balshem 2011
Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology* 2011; **64**(4):401–6. [PUBMED: 21208779]

Boneva 2001
Boneva RS, Botto LD, Moore CA, Yang Q, Correa A, Erickson JD. Mortality associated with congenital heart defects in the United States: trends and racial disparities, 1979-1997. *Circulation* 2001;**103**(19): 2376–81. [PUBMED: 11352887]

Botto 2001
Botto LD, Correa A, Erickson JD. Racial and temporal variations in the prevalence of heart defects. *Pediatrics* 2001; **107**(3):E32. [PUBMED: 11230613]

Brown 2001
Brown JW, Park HJ, Turrentine MW. Arterial switch operation: factors impacting survival in the current era. *Annals of Thoracic Surgery* 2001;**71**(6):1978–84. [PUBMED: 11426778]

Brown 2006
Brown KL, Ridout DA, Hoskote A, Verhulst L, Ricci M, Bull C. Delayed diagnosis of congenital heart disease worsens preoperative condition and outcome of surgery in neonates. *Heart (British Cardiac Society)* 2006;**92**(9): 1298–302. [DOI: 10.1136/hrt.2005.078097; PUBMED: 16449514

Bull 1999
Bull C. Current and potential impact of fetal diagnosis on prevalence and spectrum of serious congenital heart disease at term in the UK. British Paediatric Cardiac Association. *Lancet* 1999;**354**(9186):1242–7. [PUBMED: 10520632]

Carvalho 2002
Carvalho JS, Mavrides E, Shinebourne EA, Campbell S, Thilaganathan B. Improving the effectiveness of routine prenatal screening for major congenital heart defects. *Heart (British Cardiac Society)* 2002;**88**(4):387–91. [PUBMED: 12231598]

Cawsey 2016
Cawsey MJ, Noble S, Cross-Sudworth F, Ewer AK. Feasibility of pulse oximetry screening for critical congenital heart defects in homebirths. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2016;**101**(4): F349–51. [PUBMED: 26915671]

Chang 2008
Chang RK, Gurvitz M, Rodriguez S. Missed diagnosis of critical congenital heart disease. *Archives of Pediatrics & Adolescent Medicine* 2008;**162**(10):969–74. [DOI: 10.1001/archpedi.162.10.969; PUBMED: 18838650

Chew 2007
Chew C, Halliday JL, Riley MM, Penny DJ. Population-based study of antenatal detection of congenital heart disease by ultrasound examination. *Ultrasound in Obstetrics & Gynecology* 2007;**29**(6):619–24. [DOI: 10.1002/uog.4023; PUBMED: 17523161

Chu 2006
Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. Journal of Clinical Epidemiology 2006; Vol. 59, issue 12:1331-2; author reply 1332-3.

de-Wahl Granelli 2014
de-Wahl Granelli A, Meberg A, Ojala T, Steensberg J, Oskarsson G, Mellander M. Nordic pulse oximetry screening - implementation status and proposal for uniform guidelines. *Acta Paediatrica* 2014;**103**(11):1136–42. [DOI: 10.1111/apa.12758; PUBMED: 25060211

Ewer 2012b
Ewer AK. How to develop a business case to establish a neonatal pulse oximetry programme for screening of congenital heart defects. *Early Human Development* 2012; **88**(12):915–9. [DOI: 10.1016/j.earlhumdev.2012.09.013; PUBMED: 23089530

Ewer 2013
Ewer AK. Review of pulse oximetry screening for critical congenital heart defects in newborn infants. *Current Opinion in Cardiology* 2013;**28**(2):92–6. [DOI: 10.1097/HCO.0b013e32835d7e42; PUBMED: 23381095

Ewer 2014
Ewer AK. Pulse oximetry screening for critical congenital heart defects in newborn infants: should it be routine?. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2014;**99**(1):F93–5. [DOI: 10.1136/archdischild-2013-303968; PUBMED: 23934364

Ewer 2016
Ewer AK, Martin GR. Newborn pulse oximetry screening: which algorithm Is best?. *Pediatrics* 2016;**138**(5): e20161206. [DOI: 10.1542/peds.2016-1206; PUBMED: 27940777

Franklin 2002
Franklin O, Burch M, Manning N, Sleeman K, Gould S, Archer N. Prenatal diagnosis of coarctation of the aorta improves survival and reduces morbidity. *Heart* 2002;**87** (1):67–9. [PUBMED: 11751670]

Garne 2001
Garne E, Stoll C, Clementi M, Euroscan Group. Evaluation of prenatal diagnosis of congenital heart diseases by ultrasound: experience from 20 European registries. *Ultrasound in Obstetrics & Gynecology* 2001;**17**(5):386–91. [DOI: 10.1046/j.1469-0705.2001.00385.x; PUBMED: 11380961

GRADEpro GDT [Computer program]
GRADEproGDT [Computer program]. Version accessed March 2017. GRADE Handbook. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015. Available from gradepro.org.

Hall 1999
Hall DM. The role of the routine neonatal examination. BMJ (Clinical Research Ed.) 1999; Vol. 318, issue 7184: 619–20. [PUBMED: 10066183]

Han 2013
Han LM, Klewer SE, Blank KM, Seckeler MD, Barber BJ. Feasibility of pulse oximetry screening for critical congenital heart disease at 2643-foot elevation. *Pediatric Cardiology* 2013;**34**(8):1803–7. [DOI: 10.1007/s00246-013-0716-2; PUBMED: 23677390

Heron 2007
Heron MP, Smith BL. Deaths: leading causes for 2003. *National Vital Statistics Reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* 2007;**55**(10):1–92. [17408087]

Hultcrantz 2017
Hultcrantz M, Rind D, Akl EA, Treweek S, Mustafa RA, Iorio A, et al. The GRADE Working Group clarifies the construct of certainty of evidence. *Journal of Clinical Epidemiology* 2017;**87**:4–13.

Iyengar 2014
Iyengar H, Kumar P, Kumar P. Pulse-oximetry screening to detect critical congenital heart disease in the neonatal intensive care unit. *Pediatric Cardiology* 2014;**35**(3): 406–10. [DOI: 10.1007/s00246-013-0793-2; PUBMED: 24026810

Knapp 2010
Knapp AA, Metterville DR, Kemper AR, Prosser L, Perrin JM. Evidence review: critical congenital cyanotic heart disease. Prepared for: Maternal And Child Health Bureau. Massachusetts General Hospital, Center for Child and Adolescent Health Policy and Health Resources and Services Administration (HRSA), U.S. Department of Health and Human Services (DHHS) 2010. [https://www.hrsa.gov/advisorycommittees/mchbadvisory/heritabledisorders/nominatecondition/reviews/cyanoticheart.pdf]

Knowles 2005
Knowles R, Griebsch I, Dezateux C, Brown J, Bull C, Wren C. Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. *Health Technology Assessment* 2005;**9**(44):1-152, iii-iv. [PUBMED: 16297355]

Kuelling 2009
Kuelling B, Arlettaz Mieth R, Bauersfeld U, Balmer C. Pulse oximetry screening for congenital heart defects in Switzerland: most but not all maternity units screen their neonates. *Swiss Medical Weekly* 2009;**139**(47-48):699–704. [DOI: smw-12880; PUBMED: 20047132

Leeflang 2008
Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine* 2008;**149**(12):889–97. [PUBMED: 19075208]

Lloyd-Jones 2009
Lloyd-Jones D, Adams R, Carnethon M, De Simone G, Ferguson TB, Flegal K, et al. American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics - 2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 2009;**119**(3):480–6. [DOI: 10.1161/CIRCULATIONAHA.108.191259; PUBMED: 19171871

Macaskill 2010
Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. Available from srdta.cochrane.org.

Mahle 2008
Mahle WT. Physical examination and pulse oximetry in newborn infants: out with the old, in with the new?. Journal of Pediatrics 2008; Vol. 152, issue 6:747–8. [DOI: 10.1016/j.jpeds.2008.01.043; PUBMED: 18492505

Mahle 2009
Mahle WT, Newburger JW, Matherne GP, Smith FC, Hoke TR, Koppel R, et al. American Heart Association Congenital Heart Defects Committee of the Council on Cardiovascular Disease in the Young, Council on Cardiovascular Nursing, and Interdisciplinary Council on Quality of Care and Outcomes Research, American Academy of Pediatrics Section on Cardiology, Cardiac Surgery. Committee on Fetus and Newborn. Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the AHA and AAP. *Pediatrics* 2009;**124**(2):823–36. [DOI: 10.1542/peds.2009-1397; PUBMED: 19581259

Mahle 2012
Mahle WT, Martin GR, Beekman RH 3rd, Morrow WR, Section on Cardiology and Cardiac Surgery Executive Committee. Endorsement of Health and Human Services recommendation for pulse oximetry screening for critical congenital heart disease. *Pediatrics* 2012;**129**(1):190–2. [DOI: 10.1542/peds.2011-3211; PUBMED: 22201143

Manzoni 2017
Manzoni P, Martin GR, Sanchez Luna M, Mestrovic J, Simeoni U, Zimmermann LJI, et al. Pulse oximetry screening for critical congenital heart defects: a European consensus statement. *Lancet Child and Adolescent Health* 2017;**1**:88–90. [DOI: 10.1016/S2352-4642(17)30066-4

Mellander 2006
Mellander M, Sunnegardh J. Failure to diagnose critical heart malformations in newborns before discharge - an

increasing problem?. *Acta Paediatrica* 2006;**95**(4):407–13. [DOI: 10.1080/08035250500541910; PUBMED: 16720486

Narayen 2016b

Narayen IC, Blom NA, Ewer AK, Vento M, Manzoni P, te Pas AB. Aspects of pulse oximetry screening for critical congenital heart defects: when, how and why. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2016;**101**: F162-7.

Narayen 2017

Narayen IC, Kaptein AA, Hogewoning JA, Blom NA, Te Pas AB. Maternal acceptability of pulse oximetry screening at home after home birth or very early discharge. *European Journal of Pediatrics* 2017;**176**(5):669–72. [DOI: 10.1007/s00431-017-2883-2; PUBMED: 28281093

O'Donnell 2007

O'Donnell CP, Kamlin CO, Davis PG, Carlin JB, Morley CJ. Clinical assessment of infant colour at delivery. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2007;**92**(6):F465–7. [DOI: 10.1136/adc.2007.120634; PUBMED: 17613535

Office of National Statistics, 2015

Office of National Statistics (ONS). Statistical bulletin: deaths registered in England and Wales: 2014. Deaths, stillbirths and infant mortality including death rates, causes, age, and area of residence. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregistrationsummarytables/previousReleases 2015.

Peterson 2013

Peterson C, Grosse SD, Oster ME, Olney RS, Cassell CH. Cost-effectiveness of routine screening for critical congenital heart disease in US newborns. *Pediatrics* 2013;**132**(3): e595–603. [DOI: 10.1542/peds.2013-0332; PUBMED: 23918890

Powell 2013

Powell R, Pattison HM, Bhoyar A, Furmston AT, Middleton LJ, Daniels JP, et al. Pulse oximetry screening for congenital heart defects in newborn infants: an evaluation of acceptability to mothers. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2013;**98**(1):F59–63. [DOI: 10.1136/fetalneonatal-2011-301225; PUBMED: 22611113

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005; **58**(10):982–90. [DOI: 10.1016/j.jclinepi.2005.02.022; PUBMED: 16168343

RevMan 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Roberts 2012

Roberts TE, Barton PM, Auguste PE, Middleton LJ, Furmston AT, Ewer AK. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a cost-effectiveness analysis. *Archives of Disease in Childhood* 2012; **97**(3):221–6. [DOI: 10.1136/archdischild-2011-300564; PUBMED: 22247242

SAS Institute Inc. 2004

SAS Institute Inc. What's New in SAS® 9.0, 9.1, 9.1.2, and 9.1.3. SAS OnlineDoc® 9.1.3. Cary, NC 2004. [https://support.sas.com/documentation/onlinedoc/91pdf/sasdoc˙913/whatsnew˙10878.pdf]

Schunemann 2008

Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. GRADE Working Group. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ (Clinical Research Ed.)* 2008;**336** (7653):1106–10. [DOI: 10.1136/bmj.39500.677199.AE; PUBMED: 18483053

Suresh 2013

Suresh GK. Pulse oximetry screening for critical congenital heart disease in neonatal intensive care units. *Journal of Perinatology* 2013;**33**(8):586–8. [DOI: 10.1038/jp.2012.161; PUBMED: 23897309

Takwoingi 2010

Takwoingi Y, Deeks JJ. MetaDAS: A SAS macro for meta-analysis of diagnostic accuracy studies. Quick reference and worked example. Available from http://srdta.cochrane.org/ Version 1.3. Accessed July 2010.

Tegnander 2006

Tegnander E, Williams W, Johansen OJ, Blaas HG, Eik-Nes SH. Prenatal detection of heart defects in a non-selected population of 30,149 fetuses - detection rates and outcome. *Ultrasound in Obstetrics & Gynecology* 2006;**27**(3):252–65. [DOI: 10.1002/uog.2710; PUBMED: 16456842

Thangaratinam 2007

Thangaratinam S, Daniels J, Ewer AK, Zamora J, Khan KS. Accuracy of pulse oximetry in screening for congenital heart disease in asymptomatic newborns: a systematic review. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2007;**92**(3):F176–80. [DOI: 10.1136/adc.2006.107656; PUBMED: 17344253

Thangaratinam 2012

Thangaratinam S, Brown K, Zamora J, Khan KS, Ewer AK. Pulse oximetry screening for critical congenital heart defects in asymptomatic newborn babies: a systematic review and meta-analysis. *Lancet* 2012;**379**(9835):2459–64. [DOI: 10.1016/S0140-6736(12)60107-X; PUBMED: 22554860

Tworetzky 2001

Tworetzky W, McElhinney DB, Reddy VM, Brook MM, Hanley FL, Silverman NH. Improved surgical outcome after fetal diagnosis of hypoplastic left heart syndrome. *Circulation* 2001;**103**(9):1269–73. [PUBMED: 11238272]

Westin 2006

Westin M, Saltvedt S, Bergman G, Kublickas M, Almstrom H, Grunewald C, et al. Routine ultrasound examination at 12 or 18 gestational weeks for prenatal detection of major congenital heart malformations? A randomised controlled trial comprising 36,299 fetuses. *BJOG* 2006;**113** (6):675–82. [DOI: 10.1111/j.1471-0528.2006.00951.x; PUBMED: 16709210

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529–36. [DOI: 10.7326/0003-4819-155-8-201110180-00009; PUBMED: 22007046

Wren 1999

Wren C, Richmond S, Donaldson L. Presentation of congenital heart disease in infancy: implications for routine examination. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 1999;**80**(1):F49–53. [PUBMED: 10325813]

Wren 2008

Wren C, Reinhardt Z, Khawaja K. Twenty-year trends in diagnosis of life-threatening neonatal cardiovascular malformations. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2008;**93**(1):F33–5. [DOI: 10.1136/adc.2007.119032; PUBMED: 17556383

Wright 2014

Wright J, Kohn M, Niermeyer S, Rausch CM. Feasibility of critical congenital heart disease newborn screening at moderate altitude. *Pediatrics* 2014;**133**(3):e561–9. [DOI: 10.1542/peds.2013-3284; PUBMED: 24567022

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Arlettaz 2006

Study characteristics	
Patient sampling	Prospective multicenter study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: Switzerland **Setting**: 4 hospitals in Zurich: 3 maternity hospitals and the Division of Cardiology of the University Children's Hospital **Study period**: 1-year period (from May 13, 2003, to May 12, 2004) **Inclusion criteria**: all newborn infants from 35 weeks' gestation **Exclusion criteria**: Premature infants below 35 weeks' gestation Infants with a respiratory disorder Live birth cohort, n = 3663 (401 infants excluded according to exclusion criteria) **N screened**: 3262 (89%) (1764 at the University Hospital, 1011 at the Zollikerberg Hospital, 487 at the Triemli Hospital) **Gestational age**: median: 39 weeks (range 35 to 42) **Prevalence of CCHD**: 3.7 per 1000 live births
Index tests	Pulse oximetry was performed with the Nellcor NPB-40 handheld pulse oximeter and the Nellcor Max-N Oximax adhesive sensors **Screening protocol**: Site of testing: right or left foot Test timing: within 24 hours (*in 48 cases [1%], pulse oximetry was performed too early in part because of immediate postnatal transfer to the cardiology unit or because patients were discharged before 6 hours of age; in 255 cases (8%), pulse oximetry was performed after 12 hours; 2959 measurements [91%] were performed at between 6 and 12 hours*) Oxygen saturation: functional Threshold: < 95% Measurement did not exceed 2 minutes. If saturation was below 95%, a senior house officer performed a full clinical examination of the newborn. If the infant had saturation below 90% or any signs suggestive of a CHD, echocardiography was performed immediately. In the case of an asymptomatic newborn with borderline values (90% to 94%), a second measurement was performed 4 to 6 hours later
Target condition and reference standard(s)	**Target condition**: Congenital heart disease was defined as the presence of a gross structural abnormality of the heart or intrathoracic great vessels that is actually or potentially of functional significance **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography *(complete M-mode, 2-dimensional, and Doppler echocardiograms were performed either at the University Hospital with an Acuson 128XP/10 [Siemens, Erlangen, Germany] with a 7.5-mHz transducer, or at the University Children's Hospital with a Sonos 5500 [Philips, Amsterdam, Netherlands], both equipped with all Doppler modalities)* Reference standard used for negative pulse oximetry results: not stated

Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analysed: 3262)		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	**Low**
DOMAIN 4: Flow and Timing			

Arlettaz 2006 (*Continued*)

Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Unclear		
		Unclear	

Bakr 2005

Study characteristics	
Patient sampling	Prospective study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: Saudi Arabia **Setting**: neonatology department of King Abdel-Aziz Specialist Hospital **Study period**: 6-month period (January 2004 to July 2004) **Inclusion criteria**: asymptomatic newborns **Exclusion criteria**: Those admitted to the neonatal intensive care unit at birth **N screened**: 5211 **Prevalence of CCHD**: 3.7 per 1000 live births
Index tests	Pulse oximetry was performed with a Digioxi PO 920 pulse oximeter (Digicare Biomedical Technology, West Palm Beach, FL, USA) **Screening protocol**: Site of testing: right upper and lower limbs Test timing: longer than 24 hours. Average age at screening was 31.7 hours Oxygen saturation: fractional Threshold: ≤ 94% *"Any infant who had an oxygen saturation < 90% from either limb was examined by echocardiography. Saturations between 90% and 94% were verified by three readings; if they persisted in this range, echocardiography was also done."*
Target condition and reference standard(s)	**Target condition**: congenital heart disease **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: cardiology service of the only pediatric hospital in the region to identify patients who had received a diagnosis of CHD after discharge from the well-baby nursery
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n = 5211)
Comparative	

Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		

Bakr 2005 (*Continued*)

Did all patients receive a reference standard?	Unclear		
		Unclear	

Bhola 2014

Study characteristics	
Patient sampling	Retrospective observational study
Patient characteristics and setting	**Country**: Australia **Setting**: Royal Prince Alfred Hospital (tertiary maternity hospital delivering over 5000 newborns a year) **Study period**: 42-month period (from April 2008 to December 2011) **Inclusion criteria**: all newborns (routine neonatal examination) **Exclusion criteria**: not stated Live birth cohort, n = 19,765 **N screened**: 18,801 (95.1%) (648 had been admitted to the nursery and did not qualify for screening, 316 missed) **Prevalence of CCHD:** 0.2 per 1000 live births
Index tests	Pulse oximetry was performed with a Masimo Radical 5 portable oximeter (Masimo Corporation, Irvine, CA, USA) with a reusable probe with disposable Coban tape (1-inch self-adherent wrap, manufactured by 3M, Australia) **Screening protocol**: Site of testing: post-ductal (foot) Test timing: longer than 24 hours (between 24 and 72 hours of life) Oxygen saturation: functional Threshold: < 95% "*If the post-ductal saturation was 95% or more, the result was assigned as a pass. Readings between 90% and 95% led to a repeat saturation measurement in the next 1-2 hours. If the post-ductal saturation remained below 95% on repeat testing, the newborn was referred for review and examination by a senior neonatal paediatrician. If the saturation was less than 90%, at any time, the newborn was referred for review by a senior neonatal paediatrician without waiting for a repeat test.*"
Target condition and reference standard(s)	**Target condition**: congenital heart disease **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: database of the Heart Centre for Children at Children's Hospital at Westmead for any newborns undergoing cardiac surgery or catheter intervention in the first year of life
Flow and timing	Duration of follow-up: not stated Loss to follow-up: 316 missed (not performed by the resident, performed but not recorded, screening not completed before early discharge or owing to compliance issues when the new protocol was originally introduced)

Bhola 2014 (*Continued*)

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
		Unclear	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Unclear		

Bhola 2014 (*Continued*)

Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Yes		
		Unclear	

de-Wahl Granelli 2009

Study characteristics	
Patient sampling	Prospective study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: Sweden **Setting**: 5 maternity units in West Götaland **Study period**: July 2004 to March 31, 2007 **Inclusion criteria**: all newborn infants **Exclusion criteria**: Admitted to neonatal special care units Live birth cohort, n = 46,963 (7064 excluded owing to rolling start of study or admission to neonatal intensive care) Eligible, n = 39,899 (Östran n = 13,455, Mölndaln n = 8953, Trollhättan n = 7019, Borås n = 5382, Skövde n = 5090) Excluded, n = 1470 - refusal (19), oximeter failure (18), staff shortage (2), incomplete record of screening results (39) or of physical examination (1392) **N screened**: 38,429 (*flowchart page 4*) **Prevalence of CCHD**: 0.7 per 1000 live births
Index tests	Pulse oximetry was performed with a pulse oximeter Radical SET, version 4 (average time set on 8 seconds) with multisite LNOP YI sensors, Masimo, Irvine, CA, USA **Screening protocol**: Site of testing: pre-ductal (palm of right hand) and post-ductal (either foot) Test timing: longer than 24 hours Oxygen saturation: functional Threshold: < 95% "*When both pre-ductal and post-ductal oxygen saturation was < 95% or the difference between the two measurements was > 3% (≥ 2 standard deviations of interobserver measurement variability) the baby was provisionally considered to be screening positive, but a repeat measurement was performed. Babies with three repeated positive measurements were supposed to have an echocardiogram performed the same day according to the study protocol, but with some babies scheduled for early discharge only two pulse oximetry screenings were managed before the discharge examination was performed. Babies were considered screening positive until a measurement not fulfilling screening positive criteria was obtained. If saturation ≤ 90% the newborn was referred for an echocardiogram the same day.*"
Target condition and reference standard(s)	**Target condition**: congenital heart disease **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography

	Reference standard used for negative pulse oximetry results: mortality data of the National Board of Forensic Medicine (information on all deaths due to undiagnosed cardiovascular malformations in children younger than 1 year in Sweden born during the study)		
Flow and timing	Duration of follow-up: not stated Inconclusive results: 73 (results for only 1 site [34], oxygen saturation < 90% but not optimal [39]) N analyzed: 39,821		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

de-Wahl Granelli 2009 (*Continued*)

		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	No		
Was there at least 28 days of appropriate follow up?	Yes		
Did all patients receive a reference standard?	Yes		
		Low	

Ewer 2011

Study characteristics	
Patient sampling	Prospective multicenter study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: United Kingdom **Setting**: 6 obstetrical units in the West Midlands **Study period**: February 2008 to January 2009 **Inclusion criteria**: asymptomatic newborns (gestation > 34 weeks) Study includes newborns in whom congenital heart defects were suspected antenatally after midtrimester ultrasonography **Exclusion criteria**: Newborns with symptoms suggestive of cardiac disease that were detected before screening Livebirth cohort, n = 26,513 (3768 missed, 2005 declined, 685 ineligible) **N screened**: 20,055 (75.6%) **Prevalence of CCHD**: 1.2 per 1000 live births
Index tests	Pulse oximetry was performed with the Radical-7 pulse oximeter with reusable probe LNOP Y1 (Masimo, Irvine, CA, USA) **Screening protocol**: Site of testing: pre-ductal (right hand) and post-ductal (either foot in non-specified order) Test timing: within 24 hours (median age at testing of 12.4 hours for the full cohort) Oxygen saturation: functional Threshold: < 95% "*A saturation of less than 95% in either limb or a difference of more than 2% between the limb saturation readings (if both were ≥ 95%) was judged to be abnormal. Clinical examination was expedited if an abnormal test result was obtained. If this examination was unremarkable, oximetry was repeated 1to 2 hours later. If abnormalities of the cardiovascular system were detected with expedited examination, or saturations remained abnormal during a second test, the newborn were classified as test positive.*"

Target condition and reference standard(s)	**Target condition**: critical congenital heart defects (ie, death or requiring invasive intervention before 28 days) All infants with hypoplastic left heart, pulmonary atresia with intact ventricular septum, simple transposition of the great arteries, or interruption of the aortic arch All infants dying or requiring surgery within the first 28 days of life with coarctation of the aorta, aortic valve stenosis, pulmonary valve stenosis, tetralogy of Fallot, pulmonary atresia with ventricular septal defect, or total anomalous pulmonary venous connection **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical follow-up, use of cardiology databases and congenital anomaly registries
Flow and timing	Duration of follow-up: up to 12 months Loss to follow-up: none (n analyzed: 20,055)
Comparative	
Notes	Funding: National Institute for Health Research Health Technology Assessment (NIHR HTA) program (project number 06/06/03)

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			

Ewer 2011 (*Continued*)

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Yes		
Did all patients receive a reference standard?	Yes		
		Low	

Gomez-Rodriguez 2015

Study characteristics	
Patient sampling	Cross-sectional prospective study
Patient characteristics and setting	**Country**: Mexico **Setting**: Department of Neonatology, UMAE 48-Instituto Mexicano del Seguro Social (IMSS), León, Gto **Study period**: July 2010 to April 2011 **Inclusion criteria:** newborns > 6 hours of age in whom no CHD was suspected; only tested consecutive newborns who were available during the working hours of investigators **Exclusion criteria:** newborns with lung disease no informed consent **N screened**: 1037 **Gestational age**: mean (SD): 38.9 (1.1) weeks **Prevalence of CCHD**: 1.9 per 1000 live births
Index tests	Pulse oximetry was performed with a Rad-5 handheld pulse oximeter with multisite sensor **Screening protocol**: Site of testing: left lower extremity (post-ductal) Test timing: within 24 hours (mean age at pulse oximetry screening 12 hours - range 6 to 48 hours) Oxygen saturation: functional

	Threshold: < 95% Measurement was taking during 2 minutes until the reading remained the same in 2 determinations		
Target condition and reference standard(s)	**Target condition**: critical congenital heart disease (no definition included) **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical records of all follow-up at 6 months		
Flow and timing	Duration of follow-up: 6 months Loss to follow-up: none (n analyzed: 1037)		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		

Gomez-Rodriguez 2015 (*Continued*)

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Yes		
Did all patients receive a reference standard?	Yes		
		Low	

Jones 2016

Study characteristics	
Patient sampling	Retrospective observational study
Patient characteristics and setting	**Country**: United Kingdom **Setting**: neonatal intensive care unit, Northwick Park Hospital, Harrow, Middlesex (level-2 neonatal unit without on-site access to pediatric echocardiography) **Study period**: September 1, 2011, to August 31, 2013 **Inclusion criteria:** all newborns admitted to the neonatal unit during the study period **Exclusion criteria:** Antenatal diagnosis of CCHD Admitted to neonatal intensive care unit after birth Live birth cohort, n = 11,233 (973 neonatal unit admissions) **N screened**: 10,260 **Gestational age**: not stated **Prevalence of CCHD**: 0.2 per 1000 live births
Index tests	Type of pulse oximeter not stated **Screening protocol**: Site of testing: both pre-ductal and post-ductal Test timing: within 24 hours Oxygen saturation: not stated Threshold: ≤ 95% (or pre-ductal and post-ductal difference > 3%)

Target condition and reference standard(s)	**Target condition**: critical congenital heart disease defined as CHD resulting in death or requiring surgical intervention or therapeutic catheterization within the first 28 days of life **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: National Congenital Heart Disease Audit		
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 10,260)		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge	Unclear		

Jones 2016 (*Continued*)

of the results of the index tests?			
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Yes		
		Unclear	

Klausner 2017

Study characteristics	
Patient sampling	Retrospective observational study
Patient characteristics and setting	**Country**: USA **Setting**: 4 Yale-New Haven Health System hospitals in Connecticut **Study period**: January 1 and December 31, 2014 **Inclusion criteria**: all newborns delivered during the study period **Exclusion criteria**: Live-born infants who died before CCHD screening Antenatal screening Live birth cohort, n = 10,589 (171 [1.6%] underwent an echocardiogram before screening, and 98 [0.9%] were not screened; 96 were missed in error and parents refused in 2 instances) **N screened**: 10,320 **Gestational age**: 9584 (90.5%) were term (> 37 weeks) **Prevalence of CCHD**: 0 per 1000 live births
Index tests	Type of pulse oximeter not stated **Screening protocol**: Site of testing: both pre-ductal and post-ductal Test timing: longer than 24 hours Oxygen saturation: not stated Threshold: < 95%
Target condition and reference standard(s)	**Target condition**: critical congenital heart disease defined as structural defect associated with hypoxemia in the newborn period that requires surgical intervention before 1 year and, without intervention, can lead to significant morbidity and mortality **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiogram Reference standard used for negative pulse oximetry results: follow-up

Flow and timing	Of 10,316 infants with negative pulse oximetry at the time of birth, possible to review postdischarge records of only 52.1% (n = 5367)		
Comparative			
Notes	Study was supported in part by a National Heart, Lung, and Blood Institute Medical Student Research Fellowship, National Institutes of Health (award T35HL007649; to Ms Klausner), and by grant UL1 TR001863 from the National Center for Advancing Translational Science at the National Institutes of Health and the NIH Roadmap for Medical Research. Funded by the National Institutes of Health (NIH)		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

Klausner 2017 (*Continued*)

		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	No		
		Unclear	

Koppel 2003

Study characteristics	
Patient sampling	Prospective study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: USA **Setting**: well infant nurseries at 2 hospitals (New York) **Study period**: from May 1998 to November 1999 **Inclusion criteria**: all asymptomatic newborns **Exclusion criteria**: Admitted to neonatal special care units (infants who did manifest any of these clinical findings: cyanosis, tachypnea [respiratory rate: 60/min], grunting, flaring, retraction, murmur, active precordium, or diminished pulses) **N screened**: 11,281 (8642 at hospital A, 2639 at hospital B) **Prevalence of CCHD**: 0.4 per 1000 live births
Index tests	Pulse oximetry was performed with an Ohmeda Medical pulse oximeter **Screening protocol**: Site of testing: post-ductal Test timing: longer than 24 hours. Timing of oximetry determination was linked to state-mandated metabolic screening (24 hours of age) at hospital A. At hospital B, screening was performed immediately before discharge as part of a series of discharge procedures (average length of stay for vaginal delivery: 56.9 hours; for cesarean section: 103.2 hours) Oxygen saturation: functional Threshold: $\leq$ 95% "*single determination of post-ductal saturation*"
Target condition and reference standard(s)	**Target condition**: critical congenital cardiovascular malformation (CCVM) defined as a lesion that would likely require surgical correction during the first month of life **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical follow-up and New York State

Koppel 2003 (*Continued*)

	Congenital Malformations Registry (CMR)		
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed = 11,281)		
Comparative			
Notes	Study was supported by a cooperative agreement from the Centers for Disease Control and Prevention Oximeters were provided by Ohmeda Medical.		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Low	**Low**

Koppel 2003 (*Continued*)

DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Yes		
		Low	

Meberg 2008

Study characteristics	
Patient sampling	Prospective multicenter study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: Norway **Setting**: 14 hospitals with obstetrical departments and pediatric services and neonatal special or intensive care units (50% of all deliveries in Norway) **Study period**: 1-year period (2005 to 2006) **Inclusion criteria**: healthy newborns **Exclusion criteria**: Prenatal diagnosis Live birth cohort, n = 57,959 (not screened: 7951 [14%]; 224 of 7951 newborns [3%] had CHD) **N screened**: 50,008 (86%) **Prevalence of CCHD:** 0.7 per 1000 live births
Index tests	Pulse oximetry was performed with a pulse oximeter type RAD-5v (Masimo Corporation, Irvine, CA) with a multisite reusable sensor (LNOP YI) **Screening protocol**: Site of testing: post-ductal (foot) Test timing: within 24 hours (first day) Oxygen saturation: functional Threshold: < 95% "*The probe was attached for at least 2 minutes, until a stable value was obtained. Retest if the result of pulse-ox was < 95%.*"
Target condition and reference standard(s)	**Target condition**: congenital heart defects **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical follow-up
Flow and timing	Duration of follow-up: 6 months after the last infants were born Loss to follow-up: none (n analyzed: 50,008)

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
		Unclear	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		

Meberg 2008 (*Continued*)

Was there at least 28 days of appropriate follow up?	Yes		
Did all patients receive a reference standard?	Yes		
		Low	

Oakley 2015

Study characteristics	
Patient sampling	Prospective observational study
Patient characteristics and setting	**Country**: United Kingdom **Setting**: Tertiary Neonatal Unit from the Royal Gwent Hospital, Newport (3700 deliveries a year) **Study period**: 2 years (from January 2007 to December 2009) **Inclusion criteria**: all newborns at 35 weeks' gestation and above who were admitted to the postnatal ward **Exclusion criteria**: Admitted to the neonatal intensive care unit Antenatal diagnosis of CHD Births on weekends and holidays Live birth cohort, n = 9613 **N screened**: 6329 (65.8%) Gestational age of newborn infants included, range 35 to 42 weeks **Prevalence of CCHD:** 1.3 per 1000 live births
Index tests	Pulse oximetry was performed with a Nellcor NPB 40 pulse-oximeter (Pleasanton, CA) and a reusable OXI-A/N saturation probe **Screening protocol**: Site of testing: post-ductal (foot) Test timing: longer than 24 hours (all newborns were greater than 6 hours of age at the time of examination) Oxygen saturation: functional Threshold: < 95% "*Newborns with saturation readings < 95% had a repeat reading taken on the other leg after thirty minutes and if still < 95%, a further repeat reading after one hour. If the reading remained < 95%, it was considered abnormal.*"
Target condition and reference standard(s)	**Target condition**: CCHD **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: regional pediatric cardiology database and local death records
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 6369)

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
		High	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		

Oakley 2015 (*Continued*)

Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Yes		
		Low	

Ozalkaya 2016

Study characteristics	
Patient sampling	Retrospective observational study
Patient characteristics and setting	**Country**: Turkey **Setting**: Bursa Sevket Yilmaz Training and Research Hospital **Study period**: between January 2014 and December 2014 **Inclusion criteria**: asymptomatic newborns **Exclusion criteria**: Referred within first 24 hours of life or admitted to neonatal intensive care unit Perinatal CCHD Live birth cohort, n = 10,200 (excluded: hospitalized = 1100, referred = 890, perinatal CCHD = 2) **N screened**: 8208 **Gestational age**: not stated **Prevalence of CCHD**: 1 per 1000 live births
Index tests	Pulse oximetry was performed with a Nellcor pulse oximeter. **Screening protocol**: Site of testing: both pre-ductal and post-ductal Test timing: longer than 24 hours Oxygen saturation: functional Threshold: Screening test was considered positive in newborns whose saturation with pulse oximetry was less than or equal to 95% and/or who had a difference < 3% between right lower and right extremities
Target condition and reference standard(s)	**Target condition**: CCHD defined as congenital heart disease requiring catheter-based or surgical intervention within the first month of life, or causing high mortality and morbidity in the first weeks of life **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: echocardiography
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 8208)
Comparative	
Notes	

Ozalkaya 2016 (*Continued*)

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Yes		

		Low	

Richmond 2002

Study characteristics			
Patient sampling	Prospective study with consecutive enrollment of participants		
Patient characteristics and setting	**Country**: United Kingdom **Setting**: Sunderland Royal Hospital **Study period**: from April 1, 1999, to March 31, 2001 **Inclusion criteria**: asymptomatic newborn without signs of respiratory or cardiac illness **Exclusion criteria**: Admitted to neonatal care units Live birth cohort, n = 6166 (540 excluded: 447 neonatal unit, 5 no consent, 88 newborns missed) **N screened**: 5626 (91%) **Prevalence of CCHD**: 1.6 per 1000 live births		
Index tests	Pulse oximetry was performed with a radiometer Oxi machine. **Screening protocol**: Site of testing: post-ductal (foot) Test timing: within 24 hours (after the age of 2 hours and before discharge) Oxygen saturation: fractional Threshold: < 95% "*Any baby who did not achieve a post-ductal fractional saturation of at least 95% was clinically examined by the midwife. If no suspicions were raised by the examination, a second saturation measurement was performed an hour or two later. If either the examination or the repeat saturation measurement were not satisfactory, an echocardiogram was performed.*"		
Target condition and reference standard(s)	**Target condition**: congenital cardiac malformation **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: Regional Perinatal Mortality Survey and Northern Congenital Abnormality Survey & Diagnostic Database at Freeman Hospital (referral hospital)		
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 5626)		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			

Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Yes		
		Low	

Riede 2010

Study characteristics			
Patient sampling	Prospective multicenter study with consecutive enrollment of participants		
Patient characteristics and setting	**Country**: Germany **Setting**: primary, secondary, and tertiary care (34 neonatal/obstetrical departments in Saxony) **Study period**: 2-year period (from July 2006 to June 2008) **Inclusion criteria**: full-term and post-term neonates (gestational age ≥ 37 weeks) Normal routine clinical examination Informed parental consent **Exclusion criteria**: Antenatal diagnosis/suspicion of congenital heart disease Livebirth cohort, n = 48,348 (excluded: 6108 newborns [72 clinical or prenatal diagnosis of CCHD; 6036 other]) N eligible for pulse oximetry screening: 42,240 (n = 727 [91%] did not receive pulse oximetry screening, mainly because of early discharge after birth) **N screened**: 41,445 (85.7%) **Prevalence of CCHD**: 0.4 per 1000 live births		
Index tests	Pulse oximetry was performed with a great variety of devices (no further information) **Screening protocol**: Site of testing: post-ductal (foot) Test timing: longer than 24 hours Oxygen saturation: functional Threshold: ≤ 95% "*The study protocol included repeated SpO_2 measurements after 1 hour if the initial value was < 96%.*"		
Target condition and reference standard(s)	**Target condition**: CCHD **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: Saxonian perinatal and neonatal registries		
Flow and timing	Duration of follow-up: not stated Loss to follow-up: 3 for violation of study protocol (n analyzed: 41,442)		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Did the study avoid inappropriate exclusions?	No		
		Unclear	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	No		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Yes		
		Low	

Rosati 2005

Study characteristics			
Patient sampling	Prospective study with consecutive enrollment of participants		
Patient characteristics and setting	**Country**: Italy **Setting**: Perrino Hospital (referral center of the area) **Study period**: from May 1, 2000, to November 30, 2004 **Inclusion criteria:** term newborns with uncomplicated neonatal courses **Exclusion criteria**: Infants who were symptomatic (ie, heart murmur, severe cyanosis) Prenatal diagnosis of critical congenital cardiovascular malformation **N screened**: 5292 **Prevalence of CCHD**: 0.6 per 1000 live births		
Index tests	Type of pulse oximeter not stated **Screening protocol**: Site of testing: post-ductal (foot) Test timing: longer than 24 hours Oxigen saturation: functional Threshold: ≤ 95% "*Post-ductal saturation (SpO_2) and the monitoring of oxymetry values were evaluated for two minutes in each newborn.*"		
Target condition and reference standard(s)	**Target condition**: CCHD defined as lesions requiring surgical correction or interventional procedures during the first month of life **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical follow-up		
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 5292)		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
		Unclear	**Low**

Rosati 2005 (*Continued*)

DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Unclear		
		Unclear	

Sendelbach 2008

Study characteristics	
Patient sampling	Prospective study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: USA **Setting**: large public hospital (Parkland Health and Hospital System [PHHS]) in Dallas, TX, which serves a primarily indigent Hispanic population) **Study period**: from March 1, 2006, to February 28, 2007

	Inclusion criteria: Term and late preterm neonates who did not have major malformations Gestational age criteria: ≥ 35 weeks Birth weight: ≥ 2100 grams **Exclusion criteria**: Admitted to neonatal intensive care unit (NICU) Respiratory distress and/or cyanosis before 4 hours of age Live birth cohort, n = 16,432 (excluded: 66 [0.4%]; 11 had CHD) **N screened**: 15,233 (99.6%) **Prevalence of CCHD:** 0.1 per 1000 live births
Index tests	Pulse oximetry was performed with a Nellcor N-395 (Boulder, CO) pulse oximeter **Screening protocol**: Site of testing: post-ductal (foot) Test timing: within 24 hours (4 hours after delivery) Oxygen saturation: functional Threshold: < 96% "*On the day of discharge, the 4-hour pulse oximetry result was made available to the provider. A pulse oximetry result of 96% was considered normal and was not repeated. For neonates who failed to achieve 96% on the 4-hour screen, a follow-up pulse oximetry reading was performed by either the nursing staff or the medical provider by using the procedure described above. When the discharge pulse oximetry reading was < 96%, echocardiography was performed.*"
Target condition and reference standard(s)	**Target condition**: CCHD including cyanotic defects such as tetralogy of Fallot, pulmonary atresia, truncus arteriosus, transposition of the great vessels, total anomalous pulmonary venous return, and tricuspid atresia, as well as left-sided obstructive lesions, including coarctation of the aorta, critical aortic stenosis, interrupted aortic arch, and hypoplastic left heart syndrome **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical follow-up
Flow and timing	Duration of follow-up: not stated Follow-up information not available for 19 (0.1%)
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		

Sendelbach 2008 (*Continued*)

		Unclear	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	No		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	No		
		Unclear	

Singh 2014

Study characteristics	
Patient sampling	Retrospective observational study

Singh 2014 (*Continued*)

Patient characteristics and setting	**Country**: United Kingdom **Setting**: level 3 Neonatal Unit of Birmingham Women's Hospital **Study period**: from April 1, 2010, to July 31, 2013 **Inclusion criteria**: screening is part of routine practice **Exclusion criteria:** Antenatal diagnosis of CCHD **N screened**: 25,859 **Prevalence of CCHD**: 0.6 per 1000 live births
Index tests	Pulse oximetry was performed with a handheld oximeter with a reusable probe **Screening protocol**: Site of testing: post-ductal (foot) and pre-ductal (right hand) Test timing: within 24 hours Oxygen saturation: functional Threshold: < 95% "*A saturation result of < 95% in either limb or a difference of > 2% between the readings (if both were ≥ 95%) was considered abnormal. Following an abnormal first test, an initial assessment was performed. If this was unremarkable, oximetry was repeated 1to 2 hours later. If the saturations remained abnormal on second testing, or if there were concerns following the initial assessment, newborns were classified as test positive and were admitted to the neonatal unit for further assessment.*"
Target condition and reference standard(s)	**Target condition**: CCHD **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: the Regional Cardiac Centre database at Birmingham Children's Hospital, the Regional Congenital Anomaly Register, and the local mortality database
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 25,859)
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
		Unclear	**Low**

Singh 2014 (*Continued*)

DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Yes		
		Low	

Turska 2012

Study characteristics	
Patient sampling	Prospective multicenter study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: Poland **Setting**: 51 neonatal units in the Mazovian province of Poland as part of the POLKARD 2006 to 2008 program **Study period**: 1 year (from January 16, 2007, to January 31, 2008)

	Inclusion criteria: Protocol B: asymptomatic newborns at ≥ 34 weeks' gestation **Exclusion criteria**: Circulatory symptoms or coexisting diseases Prenatal diagnosis Live birth cohort, n = 55,944 (in 2611 newborns, the test could not be performed owing to technical problems [equipment failure, absence of trained staff due to holiday], 340 no consent, 1295 newborns with symptoms) **N screened**: 51,698 (92.4%) **Prevalence of CCHD**: 0.4 per 1000 live births
Index tests	Pulse oximetry was performed with Novametrix, Nellcor, and Masimo pulse oximeters **Screening protocol**: Site of testing: post-ductal (foot) Test timing: within 24 hours (between the 2nd and 24th hours of life) Oxygen saturation: functional Threshold: < 95% "*The measurement was carried out by specially trained nurses for 2 to 3 min on the infant's lower extremity between the 2nd and 24th hour of life after normalisation of the plethysmographic curve of the pulse oximeter.*"
Target condition and reference standard(s)	**Target condition**: CCHD defined as requiring an interventional procedure or cardiac surgery in the first month of life **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical follow-up or based on data from the Mazovian Centre of Public Health
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 51,698)
Comparative	
Notes	Funding: Ministry of Health in Poland

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
		Unclear	**Low**

Turska 2012 (*Continued*)

DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Unclear		
		Unclear	

Van Niekerk 2016

Study characteristics	
Patient sampling	Prospective observational study
Patient characteristics and setting	**Country**: South Africa **Setting**: Mowbray Maternity Hospital (MMH), a busy level-2 maternity hospital in the Western Cape Province, SA **Study period**: May 19 to September 19, 2014

	Inclusion criteria: All neonates > 6 hours old with no clinical signs of cardiovascular disease were eligible **Exclusion criteria**: "unwell" infants, those < 6 hours old, those born to mothers < 14 years of age or unable to give informed verbal consent (owing to illness, illiteracy, or language barriers); all infants with a prenatal diagnosis of CHD or any signs of CHD, including a heart murmur (≥ 3/6) or significant dysmorphic features Livebirth cohort, n = 2256 (1220 mothers not approached) **N screened**: 1001 (44%) **Prevalence of CCHD**: 1 per 1000 live births
Index tests	Pulse oximetry was performed with Nellcor pulse oximeters. **Screening protocol**: Site of testing: right hand and any foot Test timing: longer than 24 hours Oxygen saturation: functional Threshold: < 95%
Target condition and reference standard(s)	**Target condition**: CCHD, which leads to death or needs surgical intervention before 28 days **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: not stated
Flow and timing	Duration of follow-up: no physical follow-up Loss to follow-up: none (n analyzed: 1001)
Comparative	
Notes	Study was funded in part by the School of Child and Adolescent Health Research Committee, Department of Paediatrics, Red Cross War Memorial Children's Hospital and University of Cape Town

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
		High	**Low**
DOMAIN 2: Index Test Pulse oximetry			
If a threshold was used, was it pre-specified?	Yes		

Van Niekerk 2016 (*Continued*)

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	No		
Did all patients receive a reference standard?	No		
		Unclear	

Zhao 2014

Study characteristics	
Patient sampling	Prospective multicenter study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: China **Setting**: 18 hospitals **Study period**: from August 1, 2011, to November 30, 2012 **Inclusion criteria**: all consecutive newborns (irrespective of gestational age or neonatal intensive care unit status) **Exclusion criteria**: Prenatally diagnosed major CHD Livebirth cohort, n = 130,282 (not screened: 9575 [3571 incomplete screening data, 1450 lack of consent, 2496 transfer to superior hospital, 27 prenatally diagnosed major CHD, 2031 symptomatic

	newborns]) **N screened**: 120,707 (92.7%) **Prevalence of CCHD**: 1.2 per 1000 live births
Index tests	Pulse oximetry was performed with a RAD-5V / Multisite reusable sensor (LNOP YI, Masimo) **Screening protocol**: Site of testing: pre-ductal (right hand) and post-ductal (foot) Test timing: longer than 24 hours Oxygen saturation: functional Threshold: < 95% "*The clinician repeated pulse oximetry testing 4 hours later if the first pulse oximeter oxygen saturation measurement was between 90% and 95%. Screening was deemed positive if an SpO_2 of less than 95% was obtained both on the right hand and on either foot on two measures, separated by 4 hours; a difference between the two extremities was more than 3% on two measures, separated by 4 hours; or any measure was less than 90%.*"
Target condition and reference standard(s)	**Target condition**: CCHD **Reference standard(s)**: Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical follow-up and parents' feedback
Flow and timing	Duration of follow-up: clinical examination at 6 weeks of age at the hospital "*The Children's Hospital of Fudan University provided help with further confirmation of diagnosis for all affected babies from the participating hospitals. All cases of congenital heart disease were followed up by telephone review at least 1 year of age.*" Loss to follow-up: none (n analyzed: 120,707)
Comparative	
Notes	Funding: Key Clinical Research Project sponsored by Ministry of Health, Shanghai Public Health Three-Year Action Plan, sponsored by Shanghai Municipal Government, and National Basic Research Project of China

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
		Unclear	**Low**
DOMAIN 2: Index Test Pulse oximetry			

If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Yes		
Did all patients receive a reference standard?	Yes		
		Low	

Zuppa 2015

Study characteristics	
Patient sampling	Prospective study with consecutive enrollment of participants
Patient characteristics and setting	**Country**: Italy **Setting**: Agostino Gemelli General Hospital **Study period**: 2 years (from 2009 to 2010) **Inclusion criteria**: all newborns admitted to the nursery. These newborns by definition were considered healthy or were under observation for maternal disease, mild prematurity, or low birth weight **Exclusion criteria**:

	Newborns with syndrome **Total number of newborn infants included:** N screened: 5750 **Prevalence of CCHD:** 0.2 per 1000 births
Index tests	Pulse oximetry was performed with an Ohmeda 3900 pulse oximeter **Screening protocol**: Site of testing: post-ductal (foot) Test timing: longer than 24 hours Oxygen saturation: functional Threshold: < 95% "*The measurement was performed by a professional nurse in all newborns admitted to the nursery, between the 48th and 72nd hours of life, before discharge. The probe detector was placed on one of the two legs, making sure that the newborn was quiet and with warm ends. The measurement was performed in presence of stable, continuous and free of artefacts pulse wave, for at least 3 minutes. In case of positive screening, a second check was carried out by medical staff after 15 to 30 min.*"
Target condition and reference standard(s)	**Target condition**: CCHD defined as severe cardiac alterations that require cardiac surgery during the first year of life **Reference standard(s)**: Reference standard used for positive pulse oximetry results: electrocardiographic and echocardiography (echocardiograph "HP Sonos 4500, Agilent Technologies" [Andover, MA], a multifrequency probe [5 to 12 MHz], suitable for study of the neonatal heart. Evaluation was performed by 2-dimensional analysis [2-D], analysis of M-mode, and Doppler ultrasound) Reference standard used for negative pulse oximetry results: not stated
Flow and timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 5751)
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test Pulse oximetry			

If a threshold was used, was it pre-specified?	Yes		
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	**Low**
DOMAIN 4: Flow and Timing			
Were all patients included in the analysis?	Yes		
Was there at least 28 days of appropriate follow up?	Unclear		
Did all patients receive a reference standard?	Unclear		
		Unclear	

CCHD: critical congenital heart defect.
CHD: congenital heart defect.
NICU: neonatal intensive care unit.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Andrews 2014	Study includes only information about positive results. Investigators did not follow infants who passed the screen once they left the hospital
Ewer 2012a	Health Technology Assessment report on already included study
Hoke 2002	Case-control design
John 2016	Study includes only information about positive results.
Kardasevic 2016	Different population (not asymptomatic newborns)
Kochilas 2013	Different outcome (not accuracy)
Lhost 2014	Out-of-hospital births
Meberg 2009	Different outcome (not accuracy)
Movahedian 2016	Study includes only information about positive results.
Narayen 2016a	Out-of-hospital births
Prudhoe 2013	Study did not include enough information for construction of a 2 × 2 table. Study includes data contained in Richmond 2002 (study included)
Reich 2003	Study provides a partial 2 × 2 diagnostic table from which estimation of sensitivity was not possible
Reich 2008	Different outcome (not accuracy)
Reich 2008a	Different outcome (not accuracy)
Riede 2009	Preliminary study
Ruangritnamchai 2007	Study includes only information about positive results.
Saha 2014	Journal club
Saxena 2015	Different population (not asymptomatic newborns)
Schena 2017	Different index test (combined pulse oximetry and perfusion index)
Studer 2014	Different outcome (not accuracy)
Taksande 2013	Study includes only information about positive results. Definition of test positive was not given

(*Continued*)

Tautz 2010	No ability to determine CCHD outcomes
Tsao 2016	Study includes only information about positive results.
Vaidyanathan 2011	Study provides a partial 2 × 2 diagnostic table. No ability to determine CCHD outcomes
Valmari 2006	Case report
Walsh 2011	Different outcome (not accuracy)

CCHD: critical congenital heart defect.

DATA

Presented below are all the data for all of the tests entered into the review.

Tests. Data tables by test

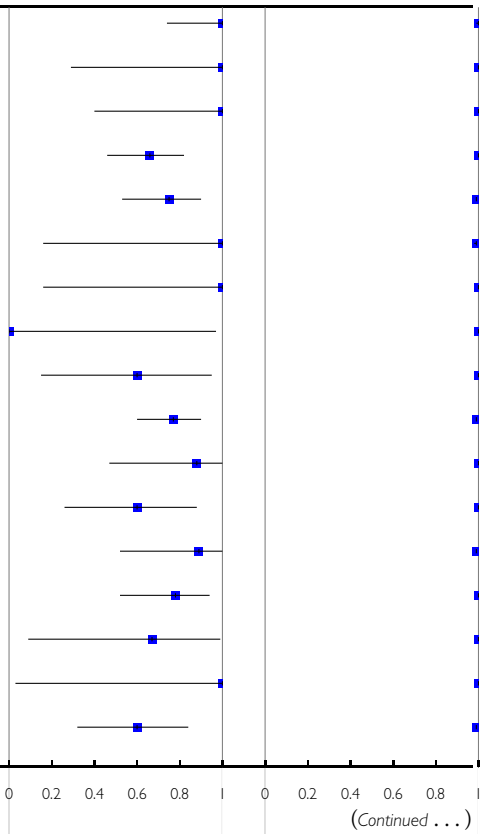
Test	No. of studies	No. of participants
1 All studies	21	457202
2 Primary analysis (threshold < 95% or ≤ 95%)	19	436758

Test 1. All studies.

Review: Pulse oximetry screening for critical congenital heart defects

Test: 1 All studies

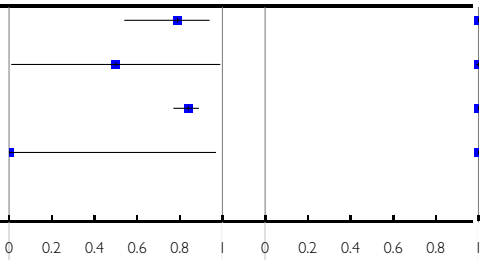
Study	TP	FP	FN	TN	Sensitivity	Specificity
Arlettaz 2006	12	12	0	3238	1.00 [0.74, 1.00]	1.00 [0.99, 1.00]
Bakr 2005	3	2	0	5206	1.00 [0.29, 1.00]	1.00 [1.00, 1.00]
Bhola 2014	4	26	0	18771	1.00 [0.40, 1.00]	1.00 [1.00, 1.00]
de-Wahl Granelli 2009	19	68	10	39724	0.66 [0.46, 0.82]	1.00 [1.00, 1.00]
Ewer 2011	18	177	6	19854	0.75 [0.53, 0.90]	0.99 [0.99, 0.99]
Gomez-Rodriguez 2015	2	12	0	1023	1.00 [0.16, 1.00]	0.99 [0.98, 0.99]
Jones 2016	2	21	0	10237	1.00 [0.16, 1.00]	1.00 [1.00, 1.00]
Klausner 2017	0	4	1	10315	0.0 [0.0, 0.97]	1.00 [1.00, 1.00]
Koppel 2003	3	1	2	11275	0.60 [0.15, 0.95]	1.00 [1.00, 1.00]
Meberg 2008	27	297	8	49676	0.77 [0.60, 0.90]	0.99 [0.99, 0.99]
Oakley 2015	7	7	1	6314	0.88 [0.47, 1.00]	1.00 [1.00, 1.00]
Ozalkaya 2016	6	1	4	8197	0.60 [0.26, 0.88]	1.00 [1.00, 1.00]
Richmond 2002	8	56	1	5561	0.89 [0.52, 1.00]	0.99 [0.99, 0.99]
Riede 2010	14	40	4	41384	0.78 [0.52, 0.94]	1.00 [1.00, 1.00]
Rosati 2005	2	1	1	5288	0.67 [0.09, 0.99]	1.00 [1.00, 1.00]
Sendelbach 2008	1	24	0	15208	1.00 [0.03, 1.00]	1.00 [1.00, 1.00]
Singh 2014	9	199	6	25645	0.60 [0.32, 0.84]	0.99 [0.99, 0.99]



(Continued . . .)

(. . . *Continued*)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Turska 2012	15	14	4	51665	0.79 [0.54, 0.94]	1.00 [1.00, 1.00]
Van Niekerk 2016	1	1	1	998	0.50 [0.01, 0.99]	1.00 [0.99, 1.00]
Zhao 2014	122	394	24	120167	0.84 [0.77, 0.89]	1.00 [1.00, 1.00]
Zuppa 2015	0	3	1	5747	0.0 [0.0, 0.97]	1.00 [1.00, 1.00]

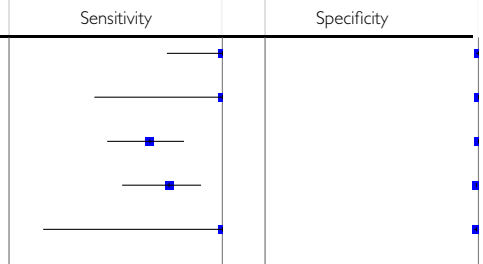


Test 2. Primary analysis (threshold < 95% or ≤ 95%).

Review: Pulse oximetry screening for critical congenital heart defects

Test: 2 Primary analysis (threshold < 95% or ≤ 95%)

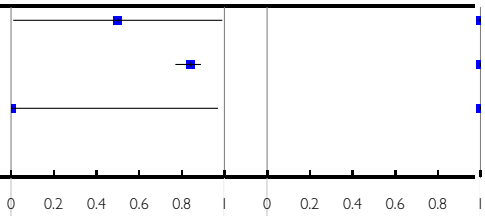
Study	TP	FP	FN	TN	Sensitivity	Specificity
Arlettaz 2006	12	12	0	3238	1.00 [0.74, 1.00]	1.00 [0.99, 1.00]
Bhola 2014	4	26	0	18771	1.00 [0.40, 1.00]	1.00 [1.00, 1.00]
de-Wahl Granelli 2009	19	68	10	39724	0.66 [0.46, 0.82]	1.00 [1.00, 1.00]
Ewer 2011	18	177	6	19854	0.75 [0.53, 0.90]	0.99 [0.99, 0.99]
Gomez-Rodriguez 2015	2	12	0	1023	1.00 [0.16, 1.00]	0.99 [0.98, 0.99]
Jones 2016	2	21	0	10237	1.00 [0.16, 1.00]	1.00 [1.00, 1.00]
Klausner 2017	0	4	1	10315	0.0 [0.0, 0.97]	1.00 [1.00, 1.00]
Koppel 2003	3	1	2	11275	0.60 [0.15, 0.95]	1.00 [1.00, 1.00]
Meberg 2008	27	297	8	49676	0.77 [0.60, 0.90]	0.99 [0.99, 0.99]
Oakley 2015	7	7	1	6314	0.88 [0.47, 1.00]	1.00 [1.00, 1.00]
Ozalkaya 2016	6	1	4	8197	0.60 [0.26, 0.88]	1.00 [1.00, 1.00]
Richmond 2002	8	56	1	5561	0.89 [0.52, 1.00]	0.99 [0.99, 0.99]
Riede 2010	14	40	4	41384	0.78 [0.52, 0.94]	1.00 [1.00, 1.00]
Rosati 2005	2	1	1	5288	0.67 [0.09, 0.99]	1.00 [1.00, 1.00]
Singh 2014	9	199	6	25645	0.60 [0.32, 0.84]	0.99 [0.99, 0.99]
Turska 2012	15	14	4	51665	0.79 [0.54, 0.94]	1.00 [1.00, 1.00]



(*Continued* . . .)

(. . . *Continued*)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Van Niekerk 2016	1	1	1	998	0.50 [0.01, 0.99]	1.00 [0.99, 1.00]
Zhao 2014	122	394	24	120167	0.84 [0.77, 0.89]	1.00 [1.00, 1.00]
Zuppa 2015	0	3	1	5747	0.0 [0.0, 0.97]	1.00 [1.00, 1.00]



ADDITIONAL TABLES

Table 1. Main studies characteristics

Study	Population	Index test					Reference standard(s)	
	Antenatal diagnosis of CHD	*Pulse oximeter*	*Limb*	*Test timing*	*Oxygen saturation*	*Threshold*	*Positive pulse oximetry*	*Negative pulse oximetry*
Arlettaz 2006	included	Nellcor NPB-40	post-ductal	within 24 hours	functional	< 95%	echocardiography	NA
Bakr 2005	excluded	Digioxi PO 920	pre-ductal and post-ductal	longer than 24 hours	fractional	≤ 94%	echocardiography	cardiology database
Bhola 2014	included	Masimo Radical 5	post-ductal	longer than 24 hours	functional	< 95%	echocardiography	cardiology database
De-Wahl 2009	excluded	Radical SET v4	pre-ductal and post-ductal	longer than 24 hours	functional	< 95%	echocardiography	mortality data
Ewer 2011	included	Radical-7	pre-ductal and post-ductal	within 24 hours	functional	< 95%	echocardiography	clinical follow-up, cardiology database & congenital registry
Gomez-Rodriguez 2015	excluded	Radical-5	post-ductal	within 24 hours	functional	< 95%	echocardiography	clinical follow-up

Table 1. Main studies characteristics (*Continued*)

Jones 2016	excluded	NA	pre-ductal and post-ductal	within 24 hours	NA	≤ 95%	echocardiography	National Congenital Heart Disease Audit
Klausner 2017	excluded	NA	pre-ductal and post-ductal	longer than 24 hours	NA	< 95%	echocardiography	clinical follow-up
Koppel 2003	excluded	Ohmeda Medical	post-ductal	longer than 24 hours	functional	≤ 95%	echocardiography	clinical follow-up & congenital registry
Meberg 2008	excluded	RAD-5v	post-ductal	within 24 hours	functional	< 95%	echocardiography	clinical follow-up
Oakley 2015	excluded	Nellcor NPB 40	post-ductal	longer than 24 hours	functional	< 95%	echocardiography	cardiology database & mortality data
Ozalkaya 2016	excluded	Nellcor	pre-ductal and post-ductal	longer than 24 hours	functional	≤ 95%	echocardiography	echocardiography
Richmond 2002	included	Oxi machine	post-ductal	within 24 hours	fractional	< 95%	echocardiography	mortality data & congenital registry
Riede 2010	excluded	NA	post-ductal	longer than 24 hours	functional	≤ 95%	echocardiography	congenital registry
Rosati 2005	excluded	NA	post-ductal	longer than 24 hours	functional	≤ 95%	echocardiography	clinical follow-up
Sendelbach 2008	excluded	Nellcor N-395	post-ductal	within 24 hours	functional	< 96%	echocardiography	clinical follow-up
Singh 2014	excluded	NA	pre-ductal and post-ductal	within 24 hours	functional	< 95%	echocardiography	mortality data & congenital registry & cardiology database

Table 1. Main studies characteristics (*Continued*)

Turska 2012	excluded	Novametrix, Nellcor & Masimo	post-ductal	within 24 hours	functional	< 95%	echocardiography	clinical follow-up and Public Health registries
Van Niekerk 2016	excluded	Nellcor	pre-ductal and post-ductal	longer than 24 hours	functional	< 95%	echocardiography	NA
Zhao 2014	excluded	RAD-5V	pre-ductal and post-ductal	longer than 24 hours	functional	< 95%	echocardiography	clinical follow-up
Zuppa 2015	excluded	Ohmeda 3900	post-ductal	longer than 24 hours	functional	< 95%	echocardiography	NA
NA: not available								

Table 2. Subgroup analysis

	N	**Sensitivity (95% CI)**	Relative sensitivity P value	**False-positive rate (FPR) (95% CI)**	Relative FPR P value
Antenatal diagnosis					
Included	4	86.3% (71.8 to 94.0)	0.071	0.46% (0.13 to 1.59)	0.231
Excluded	15	74.1% (65.7 to 81.1)		0.10% (0.05 to 0.21)	
Test timing					
Longer than 24 hours	11	73.6% (62.8 to 82.1)	0.393	0.06% (0.03 to 0.13)	0.027
Within 24 hours	8	79.5% (70.0 to 86.6)		0.42% (0.20 to 0.89)	
Limb					
Foot only	11	81.2% (70.9 to 88.4)	0.197	0.13% (0.05 to 0.31)	0.718
Foot and right hand	8	71.2% (58.5 to 81.3)		0.17% (0.06 to 0.46)	
Risk of bias ("flow and timing")					

Table 2. Subgroup analysis (*Continued*)

Unclear risk of bias	9	77.8% (64.1 to 87.3)	0.937	0.05% (0.02 to 0.12)	0.016
Low risk of bias	10	77.3% (68.8 to 84.0)		0.34% (0.17 to 0.66)	

APPENDICES

Appendix 1. Searches performed

Date: March 2017	Search strategy	Hits retrieved
Cochrane Central Register of Controlled Trials (CENTRAL; 2017, Issue 2) in the Cochrane Library	(infant or newborn or neonate or neonatal or premature or preterm or very low birth weight or low birth weight or VLBW or LBW) AND (Congenital Heart Defects OR Heart Valve Diseases OR tetralogy near fallot* OR cyanotic near heart OR congenital near heart OR congenital near cardiac OR aortic near coarctation OR valve near diseases OR hypoplastic near syndrome OR pulmonary near atresia OR interruption of the aortic arch OR valve near stenosis OR pulmonary near atresia) AND (oximetry OR pulse near oximetr* OR oxygen near saturation OR O_2 near saturation)	76
MEDLINE via PubMed (1966 to current)	(Infant, Newborn[MeSH] OR neonate* OR infant* OR newborn)* AND (Heart Defects, Congenital[MeSH] OR Heart Valve Diseases[MeSH] OR tetralogy fallot* OR cyanotic heart OR congenital heart OR congenital cardiac OR aortic coarctation OR valve diseases OR hypoplastic syndrome OR pulmonary atresia OR interruption of the aortic arch OR valve stenosis OR pulmonary atresia) AND (oximetry[MeSH] OR oximetry OR pulse oximetr* OR oxygen saturation OR O_2 saturation)	1368

(*Continued*)

Embase via Ovid (1980 to current)	(exp Infant OR exp Newborn OR neonat*.mp OR infant*.mp OR newborn*.mp) AND ((exp congenital heart malformation/) OR (exp valvular heart disease/) OR (tetralogy adj3 fallot*).mp OR (cyanotic adj3 heart).mp OR (congenital adj3 heart).mp OR (congenital adj3 cardiac).mp OR (aortic adj3 coarctation).mp OR (valve adj3 diseases).mp OR (hypoplastic adj3 syndrome).mp OR (pulmonary adj3 atresia).mp OR (interruption of the aortic arch).mp OR (valve adj3 stenosis).mp) AND (exp oximetry OR (pulse adj3 oximetr*).mp OR (oxygen adj3 saturation).mp OR (O_2 adj3 saturation).mp)	1779
CINAHL (1982 to current)	TX (Infant, Newborn OR neonate* OR infant* OR newborn*) AND TX (Congenital Heart Defects OR Heart Valve Diseases OR tetralogy fallot* OR cyanotic heart OR congenital heart OR congenital cardiac OR aortic coarctation OR valve diseases OR hypoplastic syndrome OR pulmonary atresia OR interruption of the aortic arch OR valve stenosis OR pulmonary atresia) AND TX (oximetry OR pulse oximetr* OR oxygen saturation OR O_2 saturation)	192
TOTAL before de-duplication		3415
TOTAL after de-duplication		2695

Appendix 2. QUADAS 2

Item	Criteria for assessment
Domain 1: Patient selection	
Describe methods of patient selection (prior testing, presentation, intended use of index test and setting)	
A. Risk of bias	
Was a consecutive or random sample of patients enrolled?	"Yes" if described enrolling a consecutive or random sample of newborns before discharge from hospital "No" if criteria for "yes" not achieved

(*Continued*)

	"Unclear" if the study did not describe the method of enrollment
Did the study avoid inappropriate exclusions?	"Yes" if exclusions were detailed and review authors reached consensus on the appropriateness of any exclusion "No" if inappropriate exclusions were reported (eg, if cases with antenatally diagnosed congenital heart disease were excluded) "Unclear" if insufficient information was provided
Could selection of patients have introduced bias?	A judgment of low, high, or unclear risk of bias was based on a balanced assessment of responses to the above signaling questions
B. Concerns about applicability	
Is there concern that the included patients do not match the review question?	A judgment of low, high, or unclear concern about applicability was made on the basis of how closely the sample matches an asymptomatic newborn population screened for CCHD
Domain 2: Index test	
Describe the index test and how it was conducted and interpreted	
A. Risk of bias	
Were the index test results interpreted without knowledge of results of the reference standard?	"Yes" if pulse oximetry was conducted and interpreted before the echocardiogram or clear temporal pattern to the order of testing that precludes the need for formal blinding (eg, echocardiogram, clinical follow-up, and inclusion in congenital anomaly registries are always posterior to index test) "No" if reference standard results were available to those who conducted or interpreted the pulse oximetry "Unclear" if insufficient information was provided
If a threshold was used, was it prespecified?	"Yes" if a threshold was prespecified "No" if trial authors selected a cutoff value based on analysis of collected data "Unclear" if insufficient information was provided
Could the conduct or interpretation of the index test have introduced bias?	A judgment of low, high, or unclear risk of bias was based on a balanced assessment of responses to the above signaling questions
B. Concerns about applicability	
Is there concern that the index test, its conduct, or its interpretation differ from the review question?	A judgment of low, high, or unclear concern about applicability was based on a balanced assessment of information detailed under "index test" description
Domain 3: Reference standard	

(*Continued*)

Describe the reference standard(s) and how they were conducted and interpreted	
A. Risk of bias	
Is the reference standard likely to correctly classify the target condition?	"Yes" if the study used an appropriate reference standard (diagnostic echocardiography and clinical follow-up in the first 28 days of life, including postmortem findings and mortality and congenital anomaly databases to identify false-negative patients) "No" if the study did not use an appropriate reference standard "Unclear" if the reference standard used was not clearly specified
Were the reference standard results interpreted without knowledge of results of the index test?	"Yes" if the person undertaking the reference test did not know the results of the pulse oximetry "No" if pulse oximetry results were available to those who conducted or interpreted the echocardiogram "Unclear" if insufficient information was provided
Could the reference standard, its conduct, or its interpretation have introduced bias?	A judgment of low, high, or unclear risk of bias was based on a balanced assessment of responses to the above signaling questions
B. Concerns about applicability	
Is there concern that the target condition as defined by the reference standard does not match the question?	A judgment of low, high, or unclear concern about applicability was based on the possibility of reference standards mixing both critical and non-critical congenital heart disease
Domain 4: Flow and timing	
Describe any patients who did not receive the index test and/or reference standard(s) or who were excluded from the two-by-two table (refer to flow diagram), and describe the time interval and any interventions between index test and reference standard(s)	
A. Risk of bias	
Was at least 28 days of appropriate follow-up provided?	"Yes" if follow-up was at least 28 days "No" if follow-up was less than 28 days "Unclear" if insufficient information was provided
Did all patients receive a reference standard?	"Yes" if the study specifically stated that all patients received echocardiogram, clinical follow-up, or confirmation by mortality and congenital anomaly databases (for both positive and negative pulse oximetry results) "No" if some negative pulse oximetry participants were lost to follow-up without any confirmation "Unclear" if insufficient information was provided
Were all patients included in the analysis?	"Yes" if the study had no withdrawals or withdrawals were clearly described "No" if the number of patients contributing to the two-by-two

(*Continued*)

	tables did not match the number of patients recruited and no reasons for exclusions were described "Unclear" if information was not enough to establish the flow of participants
Could the patient flow have introduced bias?	A judgment of low, high, or unclear risk of bias was based on a balanced assessment of responses to the above signaling questions

CONTRIBUTIONS OF AUTHORS

Protocol development: Shakila Thangaratinam, Andrew K. Ewer, Maria Nieves Plana, Javier Zamora, Gautham Suresh.

Selection of studies and data extraction: Maria Nieves Plana, Andrew K. Ewer.

Quality assessment: Maria Nieves Plana, Luis Fernandez-Pineda, Andrew K. Ewer.

Statistical analysis and interpretation of data: Javier Zamora, Maria Nieves Plana, Shakila Thangaratinam, Andrew K. Ewer.

Drafting of the manuscript: Maria Nieves Plana, Javier Zamora, Shakila Thangaratinam, Andrew K. Ewer.

Critical revision: Shakila Thangaratinam, Andrew K. Ewer, Maria Nieves Plana, Javier Zamora, Gautham Suresh, Luis Fernandez-Pineda.

DECLARATIONS OF INTEREST

Maria Nieves Plana: none to declare.

Javier Zamora: none to declare.

Gautham Suresh: none to declare.

Luis Fernandes-Pineda: none to declare.

Shakila Thangaratinam: an author of one of the primary studies - Ewer 2011.

Andrew K. Ewer: an author of one of the primary studies - Ewer 2011.

SOURCES OF SUPPORT

Internal sources

- IRYCIS, Spain.

Instituto Ramón y Cajal de Investigaciones Sanitarias, Madrid

- Universidad Rey Juan Carlos, Madrid, Spain.

External sources

- CIBERESP, Spain.

CIBER Epidemiology and Public Heath

- Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services, USA.

Editorial support of the Cochrane Neonatal Review Group has been funded with Federal funds from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services, USA, under Contract No. HHSN275201600005C

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Searching other resources

The protocol describes our plan to search the MEDION database (www.mediondatabase.nl), but this resource is not longer available.

Statistical analysis

We planned to use a 95% saturation level as the primary threshold for the analysis and to perform separate analyses for other thresholds categorized as < 95% and > 95% saturation levels. However, at editorial phase, it was suggested to group thresholds. As many studies used a lower than or lower than or equal to 95% threshold, we decided to group all these studies for the main analysis.

The protocol describes our plan to switch the modeling strategy to fit two univariate random-effects logistic regression models by assuming no correlation between sensitivity and specificity if the number of studies was small (fewer than four), or if the proposed modeling strategy led to problems in achieving convergence. We identified sufficient studies to fit a bivariate model and had no problem achieving model convergence.

Certainty of the evidence

We decided post hoc to assess the certainty of evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach (Schunemann 2008).

Subgroup analysis

We did not perform subgroup analyses by oxygen saturation or study design, given the low power of these subgroup analyses.

Sensitivity analysis

Additionally, we decided to perform ad hoc sensitivity analyses to describe how sensitivity and specificity vary by including or excluding studies with different thresholds.